



universität
wien

DISSERTATION

Titel der Dissertation

``Disequilibrium models, credit rationing, and bank
wholesale funding``

Verfasser

Mag. Thomas Hrdina

Angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, im Juli 2011

Studienkennzahl lt. Studienblatt:

A 094 140

Dissertationsgebiet lt. Studienblatt:

Volkswirtschaftslehre (Economics)

Betreuer:

Univ.-Prof. Dipl.-Ing. Dr. Robert M. Kunst

Univ.-Prof. Dipl.-Ing. Dr. Gerhard Sorger

Acknowledgements

When I started with the PhD programme a few years ago I only had a very broad idea about what I was engaging in, particularly with regard to writing a doctoral thesis. Recognising the scope of my decision, however, was just the first step on a long and stony road. Nevertheless, it was the right decision because I have not only acquired scientific knowledge but I also gained invaluable personal experiences. Fortunately, I was never completely on my own and therefore it is a particular concern for me to express my gratitude to those persons who accompanied me at certain stages of this venture.

First I want to thank Professor Robert M. Kunst, my main supervisor, for his many comments and suggestions from which this thesis benefited a lot. During my studies at the University of Vienna he also provided me with the econometric munition and the basic knowledge respectively I needed for writing this piece of work.

Furthermore, I am indebted to my second supervisor Professor Gerhard Sorger who was very helpful at various stages of my doctoral studies. In particular, without his encouragement and the (financial) support of the Department of Economics I would have not been able to complete my dissertation in this relatively short time period.

Finally, I am very grateful to my spouse Elisabeth for being at my side throughout this venture. For a better part, it is due to her that I stayed the course.

Contents

Acknowledgments	3
1 Introduction	8
2 Single-Market Disequilibrium Models: Insights from Monte Carlo Simulations	12
2.1 Introduction	12
2.2 Disequilibrium Models: A Short Review	13
2.2.1 Canonical Model Specifications	14
2.2.2 Maximum Likelihood Estimation Approach	15
2.2.3 Computational Issues with ML	18
2.3 Disequilibrium Model Misspecification	19
2.3.1 Design of Misspecification Experiments	21
2.3.2 Results of Misspecification Experiments	23
2.4 Dynamic Models	25
2.4.1 Design of Dynamic Model Experiments	26
2.4.2 Results of Dynamic Model Experiments	28
2.5 Conclusion	29
2.6 Appendix	30
3 Non-Price Credit Rationing by U.S. Commercial Banks During the Great Recession	35
3.1 Introduction	35
3.2 Literature Review	36
3.2.1 Indirect Approaches to Credit Rationing	38
3.2.2 Disequilibrium Econometric Models	39
3.2.3 Credit Rationing and Credit Crunches	42
3.3 Empirical Analysis	44

3.3.1	Preliminary Discussion	44
3.3.2	Models with Sample Separation Information	46
3.3.3	Demand & Supply Specification	48
3.3.4	Data & Preliminary Analysis	49
3.3.5	Estimation Results	55
3.4	Conclusion	61
3.5	Appendix	62
3.5.1	Data Description	64
3.5.2	Graphical Data Representation	65
4	Wholesale Funding and Commercial Bank Lending – Impli-	
	cations from the 2008 Liquidity Crunch	68
4.1	Introduction	68
4.2	Review	70
4.2.1	Bank Wholesale Funding	70
4.2.2	The Liquidity Crunch	72
4.2.3	Wholesale Funding, Liquidity and Bank Lending	75
4.3	Empirical Analysis	78
4.3.1	Variables and Specification	79
4.3.2	Data	82
4.3.3	Estimation Results	83
4.4	Conclusion	87
4.5	Appendix	88
	Bibliography	90
	Abstract	95
	Zusammenfassung	97
	Curriculum Vitae	99

Chapter 1

Introduction

The three chapters which form the main body of this thesis are linked methodologically as well as thematically. The predominant theme is the financial crisis that started in 2007 in the U.S. housing market. The repercussions of this crisis, which eventually evolved into a global economic crisis, are felt even today. Naturally, an event of this dimension raises a myriad of questions. For instance, how is it possible that the crisis, which was thought to be confined to a relatively small sector of the U.S. financial market, spilled over into the real economy and negatively affected economic growth around the world? These and similar questions are related to the causes and underlying mechanisms of the crisis. In order to prevent future crises of this magnitude, however, one also has to analyse potential policy and regulatory responses respectively. The thesis at hand is intended to contribute to both of these research directions.

Beyond that, chapter two and three respectively highlight a promising alternative to the equilibrium approach for the estimation of market demand and supply, namely disequilibrium econometric models. What distinguishes these models from the equilibrium framework is that they do not assume that market demand and supply are necessarily equal. Instead, the minimum of demand and supply determines the quantity transacted. Unfortunately, relaxing the equilibrium assumption generally has a number of unfavourable implications for the estimation of disequilibrium models. Most notably, the ease of application of most of these models is lower which is probably why the great majority of studies apply the equilibrium model. Two important issues in this respect are disequilibrium model specification and dynamic model features, which are the topic of the second chapter.

With regard to the former, researchers not only have to evaluate whether they should apply the equilibrium or the disequilibrium framework but also which of the various disequilibrium models fits their data best. The main distinction in this respect concerns the specification of the price adjustment process. While some models are based on the assumption that the market price is fully rigid, others assume that the price adjusts to imbalances between demand and supply, although it does not have to offset them. Most studies in the applied literature, however, do not pay attention to the issue of specification. Against this background, the first part of our analysis in chapter two deals with disequilibrium model misspecification. In particular, we evaluate the degree of misspecification of the different canonical disequilibrium models by means of Monte Carlo sampling experiments. The second part of the chapter then focuses on dynamic disequilibrium models. The main question in this respect is how important is the provision for dynamic model features since static disequilibrium estimators have been found to be consistent even if the data generating process is dynamic. The results from chapter two are not only intended to deepen our understanding of the canonical models but also to guide our analysis in chapter three in which we apply the disequilibrium framework to real world data.

The third chapter deals with the popular conjecture that during the recent financial crisis banks in the U.S. resorted to non-price credit rationing – a phenomenon which is not to be confounded with conventional price rationing. According to various accounts banks increased quantity rationing due to the turmoil in financial markets. This so-called credit crunch allegedly led to a contraction in economic activity, i.e. it is assumed to be a major transmission channel of the crisis from the financial to the real sector. Our analysis employs the disequilibrium econometric framework in order to estimate excess credit demand and the extent of non-price rationing respectively. With this estimate we then test the credit crunch hypothesis. The main innovation in this respect is the application of an excess demand indicator other than the change in the loan rate which facilitates estimation immensely. For instance, using this alternative indicator not only allows us to estimate a wide array of different disequilibrium models but also to take into account model dynamics. Chapter three also wants to raise the awareness that disequilibrium models are an important alternative to the equilibrium framework which, given the existence of an appropriate excess demand indicator, are not necessarily associated with a lower ease of application. Furthermore, our analysis has an important policy dimension. This is because the appropriate

policy measures in case of non-price credit rationing are different from the case of conventional price rationing.

Finally, the fourth chapter is concerned with bank wholesale funding and commercial bank lending. Although the econometric methodology we use in this chapter is completely different from the two previous chapters, there is a close thematic relation between the third and the fourth chapter. While chapter three analyses lending behaviour on an aggregate level, chapter four looks at the micro level, i.e. the lending behaviour of individual banks. In other words, the analyses in these two chapters complement each other. The main focus of attention in the fourth chapter is wholesale funds dependence and subsequent bank lending. Empirical evidence indicates that banks have used more and more wholesale funds to finance the asset side of their balance sheets. The turmoil in wholesale funding markets during the financial crisis implied that banks with a higher dependence on this type of funds found it increasingly difficult to secure funding liquidity in these markets. In order to meet their short-term obligations, banks therefore had to deleverage, meaning that they had to shrink the asset side of their balance sheets. This, in turn, implied that banks had to sell securities and/or stop lending. Given this stylised description we test two related hypotheses. First, whether banks with a higher wholesale dependence decreased subsequent lending more and second whether the effect of wholesale dependence on subsequent lending is a decreasing function of the liquidity of a bank's balance sheet. While some positive evidence with regard to the first hypothesis has been found, the second hypothesis has never been tested before. The results from our analysis are particularly interesting with regard to banking regulation since empirical evidence in favour of our two hypotheses would support the introduction of liquidity standards as proposed by the Basel Committee on Banking Supervision in response to the recent crisis.

Chapter 2

Single-Market Disequilibrium Models: Insights from Monte Carlo Simulations

2.1 Introduction

This chapter deals with the estimation of single-market disequilibrium models in which the price does not fully adjust to clear the market and the shorter side of the market determines the quantity transacted.¹ In particular, our analysis, which is based on Monte Carlo sampling experiments, is concerned with the consequences of disequilibrium model misspecification and dynamic models.

With regard to the former we evaluate whether the degree of misspecification increases with the size of a particular model parameter. For instance, it has been shown that there exists a close relation between disequilibrium specifications in which the price is not fully rigid, i.e. the price adjusts somewhat to imbalances between demand and supply, and the equilibrium specification respectively. This relation is based on the speed of price adjustment, denoted with γ . Most notably, one can show that for $\gamma \rightarrow \infty$ in the limit the probability density functions (p.d.f.) of the endogenous variables in these different

¹For an introduction to multi-market disequilibrium models see, for instance, Quandt (1988) or Rudebusch (1987). Also note that many authors argue that the term “disequilibrium” is a misnomer since it suggest that one faces a situation out of equilibrium. Excess demand and excess supply respectively could, however, represent equilibrium situations, albeit not in a Walrasian sense.

specifications are equivalent. This implies, however, that in terms of endogenous price models the degree of misspecification is a decreasing function of the parameter γ . While the case of a very high speed of price adjustment has been analysed thoroughly, the case of $\gamma \rightarrow 0$, corresponding to less and less price adjustment, has been neglected. This is a shortcoming because a number of empirical studies provide evidence that the speed of adjustment can be very low. We address this case and its implications in the first part of this chapter.

In the second part we discuss dynamic models. The main issue in this respect is that for disequilibrium models with unknown sample separation, i.e. models in which it is not possible to infer whether the quantity observed has been determined by the supply or the demand equation, one cannot derive the maximum likelihood estimator if the underlying model exhibits dynamic features such as autocorrelated errors and/or lagged dependent variables. This is because the likelihood function becomes intractable in these cases. While a number of authors have suggested to use simulation-based estimation methods instead, it has been shown that under certain conditions the static estimators are consistent even in a dynamic setting. Against this background, we analyse the properties of the static estimators if the true models are dynamic.

The remainder of this chapter is structured as follows. In section 2.2 we provide a short review of the disequilibrium literature where we concentrate on the canonical model specifications, the maximum likelihood estimation method and problems associated with the latter. Section 2.3 addresses the issue of disequilibrium model misspecification, while dynamic models are covered in section 2.4. Both sections start with a motivation and then give a detailed description of the Monte Carlo experiments conducted and their results. Finally, section 2.5 concludes.

2.2 Disequilibrium Models: A Short Review

The seminal article by Fair & Jaffee (1972) on the estimation of disequilibrium models was followed by a broad discussion about different model specifications, estimation methods and other, related issues. We do not review the whole disequilibrium literature here and instead concentrate on those specifications and estimation methods that are used in subsequent sections. For an in-depth treatment of the topic the reader is referred to the excel-

lent books by Quandt (1988), from which we adopt the taxonomy for the following description, and Srivastava & Rao (1990).

2.2.1 Canonical Model Specifications

The most basic disequilibrium econometric model, which is referred to as model A, consists of the following system of equations:

$$d_t = \alpha_1 p_t + \boldsymbol{\alpha}' \mathbf{x}_{d,t} + \epsilon_{d,t} \quad (2.1)$$

$$s_t = \beta_1 p_t + \boldsymbol{\beta}' \mathbf{x}_{s,t} + \epsilon_{s,t} \quad (2.2)$$

$$q_t = \min(d_t, s_t) \quad (2.3)$$

where d_t denotes (unobserved) demand in period t , s_t is (unobserved) supply, q_t is the actual (observed) quantity transacted in the market, p_t is the market price associated with the parameters α_1 and β_1 respectively, $\mathbf{x}_{d,t}$ and $\mathbf{x}_{s,t}$ are vectors of exogenous variables, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are parameter vectors and $\epsilon_{d,t}$ and $\epsilon_{s,t}$ are disturbance terms. The latter are assumed to be jointly normally distributed as well as serially and mutually independent.

The most distinguishing feature of model A – and all other disequilibrium models presented below – is that the equilibrium condition $q_t = d_t = s_t$ is replaced by equation (2.3) which implies that the actual quantity transacted in the market is the minimum of demand and supply.² Thus, it is assumed that the price does not clear the market and therefore the latter is characterised by excess demand/supply. In model A the price is even assumed to be exogenous, i.e. the price is fully rigid meaning that it does not adjust to imbalances between demand and supply. The minimum condition is usually justified on the grounds that most markets are characterised by voluntary exchange, that is neither side of the market is forced to trade more than it wishes.

Adding a deterministic price adjustment equation such as (2.4) to model A yields another disequilibrium model which is referred to as model C.³

²The minimum condition serves the same purpose as the equilibrium condition: it relates the unobserved endogenous variables d_t and s_t to the observed endogenous variable q_t . Due to the unobserved endogenous variables, disequilibrium models also belong to the class of latent variable models.

³In the literature, model C is also referred to as the quantitative model, since the change in price is proportional to the quantity of excess demand.

$$p_t = p_{t-1} + \gamma(d_t - s_t) \quad (2.4)$$

Unlike in A, the price in period t is now endogenous and a function of the quantity of excess demand – notice that different lag structures are conceivable with regard to the adjustment equation.⁴ In contrast to model A the price adjusts to market imbalances, though it does not necessarily offset them completely. Furthermore, economic theory suggests that the parameter γ , associated with the speed of price adjustment, is positive since in case of excess demand the price is likely to increase while in case of excess supply it presumably decreases – a generalisation would allow for different adjustment speeds γ_1 and γ_2 in case of excess demand and excess supply respectively. Also note that the two polar cases $\gamma = 0$ and $\gamma = \infty$ correspond to no price adjustment and perfect price adjustment – the latter means that the price adjusts to the equilibrium price immediately. If the adjustment equation is stochastic as in equation (2.5) one finally gets model D.

$$p_t = p_{t-1} + \gamma(d_t - s_t) + \epsilon_{p,t} \quad (2.5)$$

Notice that the stochastic price adjustment equation could be extended to include additional explanatory variables. As in model C, the price is endogenous and the disturbance terms are assumed to be joint normal as well as serially and mutually independent, i.e. $\boldsymbol{\epsilon}'_t = (\epsilon_{d,t} \ \epsilon_{s,t} \ \epsilon_{p,t}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ with the diagonal variance-covariance matrix $\boldsymbol{\Omega} = \text{diag}(\sigma_d^2, \sigma_s^2, \sigma_p^2)$.

2.2.2 Maximum Likelihood Estimation Approach

In the disequilibrium literature, various methods have been suggested for the estimation of the three canonical models presented above. An important observation in this respect is that the scope of available methods critically depends on the model specification and the information contained in it. Information here means knowledge about the sample separation: the minimum condition separates the observations for q_t into observations which are determined by the demand equation and observations which are determined by

⁴Maddala (1983) provides a discussion regarding different lag structures. He also offers the following interpretation of equation (2.4): the price does not rise in response to excess demand but the market is characterised by excess demand since the price does not fully adjust.

the supply equation.⁵ While from the deterministic price adjustment equation in model C one can infer the sample separation – note that the sign of $\Delta p_t = p_t - p_{t-1}$ is observed and thus the sign of $(d_t - s_t)$ and the minimum are known – there is no such knowledge either in model A or D. This lack of information, in turn, constrains the methods potentially available for estimation. Model C, for instance, can be estimated with the ordinary two-stage least squares method, while model A and D cannot – except in cases where additional information is provided about the sample separation.⁶

In subsequent sections we concentrate on the maximum likelihood (ML) method of estimation which is available for all three models. Although the likelihood functions have already been derived elsewhere, it is a worthwhile exercise to go through the derivation because it warrants a deeper insight into the models.

Sample Separation Known

With regard to model C the likelihood function is derived in the following way. First note that in case of $d_t < s_t$ the price adjustment equation (2.4) can be written as $q_t = s_t + \Delta p_t/\gamma$, while if $d_t \geq s_t$ it can be written as $q_t = d_t - \Delta p_t/\gamma$. Thus, when $d_t < s_t$ model C is given by

$$\begin{aligned} q_t &= \alpha_1 p_t + \boldsymbol{\alpha}' \mathbf{x}_{d,t} + \epsilon_{d,t} \\ q_t &= \beta_1 p_t + \boldsymbol{\beta}' \mathbf{x}_{s,t} + \Delta p_t/\gamma + \epsilon_{s,t} \end{aligned} \quad (2.6)$$

while for $d_t \geq s_t$ it is

$$\begin{aligned} q_t &= \alpha_1 p_t + \boldsymbol{\alpha}' \mathbf{x}_{d,t} - \Delta p_t/\gamma + \epsilon_{d,t} \\ q_t &= \beta_1 p_t + \boldsymbol{\beta}' \mathbf{x}_{s,t} + \epsilon_{s,t} \end{aligned} \quad (2.7)$$

Now define the two variables Δp_t^+ and Δp_t^- as follows:

⁵Because of these two regimes disequilibrium models are also referred to as switching regression models.

⁶See, for instance, Rudebusch (1987). Also note that the price in model C is a special case of an excess demand indicator.

$$\begin{aligned}\Delta p_t^+ &= \begin{cases} \Delta p_t & \text{if } \Delta p_t > 0 \\ 0 & \text{otherwise} \end{cases} \\ \Delta p_t^- &= \begin{cases} -\Delta p_t & \text{if } \Delta p_t < 0 \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

The two systems (2.6) and (2.7) can then be combined to yield⁷

$$q_t = \alpha_1 p_t + \boldsymbol{\alpha}' \mathbf{x}_{d,t} - \Delta p_t^+ / \gamma + \epsilon_{d,t} \quad (2.8)$$

$$q_t = \beta_1 p_t + \boldsymbol{\beta}' \mathbf{x}_{s,t} - \Delta p_t^- / \gamma + \epsilon_{s,t} \quad (2.9)$$

Finally, since the joint density of $\epsilon_{d,t}$ and $\epsilon_{s,t}$ is known, one can infer the log-likelihood function for model C which is given by

$$\log L_c = T \log(|\beta_1 - \alpha_1 + 1/\gamma|) - T \log(2\pi\sigma_d\sigma_s) - \sum_{t=1}^T \left[-\frac{1}{2} \left(\frac{\epsilon_{d,t}^2}{\sigma_d^2} - \frac{\epsilon_{s,t}^2}{\sigma_s^2} \right) \right]$$

Sample Separation Unknown

As one can see from above, the derivation of the log-likelihood function for model C is based on the knowledge about sample separation. In model A and D respectively, in which sample separation is unknown, the procedure is necessarily different. Consider model D, for instance – the derivation of the log-likelihood function for model A is analogous, except that the price is exogenous. In model D the joint density of d_t , s_t and p_t , which can be derived from the joint density of $\epsilon_{d,t}$, $\epsilon_{s,t}$ and $\epsilon_{p,t}$, is given by

$$g(d_t, s_t, p_t) = \frac{|J|}{(2\pi)^{3/2} |\boldsymbol{\Omega}|^{1/2}} \exp \left[-(1/2) \boldsymbol{\epsilon}_t' \boldsymbol{\Omega}^{-1} \boldsymbol{\epsilon}_t \right]$$

where $|J|$ is the absolute value of the Jacobian determinant from the transformation of the disturbances to d_t , s_t and p_t , $\boldsymbol{\Omega}$ is the variance-covariance matrix of the disturbance terms and

⁷Notice that equation (2.8) and (2.9) could also be estimated with two-stage least squares.

$$\boldsymbol{\epsilon}_t = \begin{pmatrix} \epsilon_{d,t} \\ \epsilon_{s,t} \\ \epsilon_{p,t} \end{pmatrix} = \begin{pmatrix} d_t - \alpha_1 p_t - \boldsymbol{\alpha}' \mathbf{x}_{d,t} \\ s_t - \beta_1 p_t - \boldsymbol{\beta}' \mathbf{x}_{s,t} \\ p_t - p_{t-1} - \gamma(d_t - s_t) \end{pmatrix}$$

Given this density, the next step in the derivation of the log-likelihood function is to state the model in terms of the observable endogenous variables q_t and p_t . First note that the probability that any observation q_t belongs to the demand curve is equal to $\lambda_t = \text{Prob}(d_t < s_t)$. The probability that any observation q_t belongs to the supply curve is then $(1 - \lambda_t)$. Thus, the conditional joint density of q_t and p_t for the case that $q_t = d_t$ and $s_t > q_t$ is given by

$$h(q_t, p_t | q_t = d_t) = \frac{\int_{q_t}^{\infty} g(q_t, s_t, p_t) ds_t}{\lambda_t}$$

– note that the denominator λ_t , i.e. the probability that q_t belongs to the demand curve, is equal to the numerator integrated over q_t . Similarly, the conditional joint density of q_t and p_t for the case that $q_t = s_t$ and $d_t \geq q_t$ is given by

$$h(q_t, p_t | q_t = s_t) = \frac{\int_{q_t}^{\infty} g(d_t, q_t, p_t) dd_t}{1 - \lambda_t}$$

Finally, the unconditional joint density of q_t and p_t can be stated as

$$\begin{aligned} h(q_t, p_t) &= \lambda_t h(q_t, p_t | q_t = d_t) + (1 - \lambda_t) h(q_t, p_t | q_t = s_t) \\ &= \int_{q_t}^{\infty} g(q_t, s_t, p_t) ds_t + \int_{q_t}^{\infty} g(d_t, q_t, p_t) dd_t \end{aligned}$$

Due to the assumption that the disturbances are serially independent, the log-likelihood function is then given by

$$\log L_D = \sum_{t=1}^T \log(h(q_t, p_t))$$

2.2.3 Computational Issues with ML

A drawback of the maximum likelihood approach is that even in cases where the first and second order conditions for the maximisation of the likelihood

function can be derived analytically, one usually has to employ numerical optimization algorithms – like the Marquardt or the BHHH procedure – since these conditions are complicated non-linear functions. Nevertheless, if the likelihood function is well-behaved and bounded from above the iterative optimisation algorithms have a good chance of converging to a maximum and the choice of the starting values should not be of vital importance. Unfortunately, though, it is a well-known fact that the likelihood functions in model A and D respectively are ill-behaved. Besides the possibility of multiple maxima, one can show that for certain parameter values the functions are unbounded which makes maximisation extremely difficult.⁸ The unboundedness problem is a result of the substantial latency in the two models: not only are the quantities demanded and supplied unobserved but sample separation is also unknown – notice that unboundedness is not a problem in model C. According to Quandt (1988), another and probably even more severe issue with respect to the estimation of disequilibrium models is the problem of false or spurious maxima. This problem occurs if the iterative optimisation algorithm drives the estimates into a direction where the likelihood function is not defined anymore – the latter happens, for instance, if the covariances between the disturbance terms are non-zero and the variance-covariance matrix becomes singular. In such cases computation usually breaks down at estimates which are far from the true parameter values.

The computational problems outlined above are a major concern. Our main focus of attention in the following two sections, however, is directed towards the issue of disequilibrium model misspecification and dynamic models.

2.3 Disequilibrium Model Misspecification

In estimating market demand and supply the researcher faces the problem of choosing the most appropriate model specification for the data at hand. To facilitate her choice, a number of tests have been developed. One set of tests is concerned with whether to choose the standard equilibrium model or a disequilibrium model. Hwang (1980) introduced a simple test to distinguish between the equilibrium and the disequilibrium hypothesis (model

⁸More on this point is provided in Maddala (1983), for instance. Quandt (1988) notes in this respect that in order to circumvent the unboundedness problem one should avoid the maximum likelihood method and use non-linear least squares instead. Laroque & Salanie (1994), however, find that the performance of the latter is very poor.

A). Quandt (1988), in turn, presents test procedures with which one can distinguish between the equilibrium model and model C and D respectively. A second group of misspecification tests, in contrast, focuses on the different disequilibrium specifications. A Lagrange multiplier test by Hajivassiliou (1986), for example, tests for price endogeneity and supports the researcher in determining whether model A or a model in which the price is endogenous, as in model C or D, is more appropriate. Goldfeld & Quandt (1981), on the other hand, present a simple procedure to test whether model C or model D should be selected.

In addition, they analyse the consequences of estimating a misspecified endogenous price model, e.g. estimating model C although the data have been generated by model D or the equilibrium model. In doing so, they emphasise the asymptotic relation between model C and D respectively on the one hand and the equilibrium model on the other, namely that for $\gamma \rightarrow \infty$ the p.d.f. in model C and D converge to the p.d.f. in the equilibrium model, implying that in the limit the estimators of these models are the same. In other words the equilibrium model is nested both in model C and D. Given this result, Goldfeld & Quandt then show by means of sampling experiments that as γ increases the degree of misspecification of model C, D and the equilibrium model respectively decreases.

Our analysis, in contrast, concentrates on cases where the speed of price adjustment is very low, i.e. in which $\gamma \rightarrow 0$, and the polar case of $\gamma = 0$ respectively. Given the dominance of the equilibrium model it is not astonishing that most studies are concerned with the opposite case. However, various applied studies, such as Laffont & Garcia (1977), Sealey (1979) or Kugler (1987), find that in many markets the speed of price adjustment is sluggish at best. Furthermore, the case of very slow price adjustment is closely related to disequilibrium model A, in which the price is fully rigid, and therefore our analysis sheds light on the degree of misspecification among the different canonical disequilibrium models. For instance, we argue below that under the true model C or D the degree of misspecification of model A – reflected in the quality of the estimates it provides – should decrease for lower true values of γ . And finally, since model A has recently enjoyed great popularity in the applied disequilibrium literature, our results have strong implications for the validity of these studies.⁹

⁹Examples in this respect are Pazarbasioglu (1997), Gosh & Gosh (1999), Kim (1999), Barajas & Steiner (2002), Nehls & Schmidt (2003), Ikhida (2003) and Allain & Oulidi

Based on a preliminary examination of the different model specifications we expect to find the following results. First, it is straightforward to see that if $\gamma \rightarrow 0$ either in model C or D then in the limit the price becomes exogenous in both models. In particular, for $\gamma = 0$ the non-stochastic price adjustment equation (2.4) implies that the price is constant while in case of the stochastic price adjustment equation (2.5) the price follows a random walk. This implies that two special variants of model A are nested in model C and D respectively.¹⁰ For this reason we anticipate that the degree of misspecification of model A is an increasing function of γ both under the true model C and D.¹¹ Finally, if the true model is A and the speed of price adjustment is implicitly equal to zero, we believe that the performance of both estimator C and D are highly sensitive to the starting value provided for γ : for high values estimator C and D should perform worse than for low values – in the following we use the term performance to describe the quality or accuracy of the estimates provided by the various disequilibrium estimators.

In order to test these hypotheses and to highlight the consequences of disequilibrium model misspecification, particularly with regard to model A, we have conducted a number of Monte Carlo sampling experiments which are presented in the following section.

2.3.1 Design of Misspecification Experiments

Table 2.1 in the appendix provides an overview of the Monte Carlo misspecification experiments we have conducted. As one can see from the table, there are four groups of experiments. In the first group we analyse the performance of the equilibrium estimator and estimator A in case the data have been generated by the equilibrium model and model A respectively. We do not expect estimator A to perform well under the true equilibrium model and vice versa. Next, in experiment 2.1 to 2.3 the true model is A and we test the performance of estimator C. These experiments differ from each

(2009).

¹⁰Note that while in the polar case of $\gamma = 0$ estimator A is closely related to the estimators of model C and D, this relation is somewhat looser than between the equilibrium estimator and estimators C and D in the polar case of $\gamma = \infty$.

¹¹In a seminal contribution Maddala & Nelson (1974) have stumbled across this result before. They show by means of a small-scale Monte Carlo simulation that estimator A performs well even though the data have been generated by a model very similar to C.

other in the choice of the starting value for γ . The other experiments in this group, i.e. experiment 2.4 to 2.6, are designed in a similar fashion, with the only difference that they test the performance of estimator D. Finally, in the third and fourth group of experiments we test the performance of estimator A under the true model C and D respectively. As indicated above, we expect estimator A to perform better the lower the true value of γ . Also note that in contrast to the first two groups of experiments we estimate model A, C and D in group three and four. This is because of the close relation between specification C and D: for $\sigma_p^2 \rightarrow 0$ the p.d.f. in model D converges to the p.d.f. in model C.

The number of replications S and the number of observations T is the same for all experiments. The choice of $S = 500$ reflects our interest to achieve both a high accuracy in terms of our results and a relatively short computing time while $T = 100$ is typically encountered in applied work, e.g. with quarterly data. The models which we used to generate our artificial data can be summarised as follows:

$$\begin{aligned} d_t &= \alpha_1 + \alpha_2 x_{1,t} + \alpha_3 p_t + \epsilon_{d,t} \\ s_t &= \beta_1 + \beta_2 x_{2,t} + \beta_3 p_t + \epsilon_{s,t} \\ q_t &= \begin{cases} d_t = s_t & \text{for the Equilibrium Model} \\ \min(d_t, s_t) & \text{for Model A/C/D} \end{cases} \\ p_t &= \begin{cases} p_{t-1} + \gamma(d_t - s_t) & \text{for Model C} \\ p_{t-1} + \gamma(d_t - s_t) + \epsilon_{p,t} & \text{for Model D} \end{cases} \end{aligned}$$

where $\alpha_1 = 2$, $\alpha_2 = 1$, $\alpha_3 = -0.5$, $\beta_1 = 1.4$, $\beta_2 = 0.8$ and $\beta_3 = 0.1$. The disturbance terms $\epsilon_{d,t}$, $\epsilon_{s,t}$ and $\epsilon_{p,t}$, which are randomly drawn in each replication, have a normal distribution with mean zero and variance $\sigma_d^2 = 1.44$, $\sigma_s^2 = 1$ and σ_p^2 . The two exogenous variables x_1 and x_2 , which are fixed in every replication, have been generated from a uniform distribution over the range (10,15) and (11,17) respectively. In addition, for the experiments in group three and four we need an observation on the initial price level for which we choose $p_0 = 3.5$. For experiments where the true model is A, we assume that the exogenous price series is drawn from a uniform distribution over the range (2,5) and is fixed for all replications. The true parameter values for γ and σ_p^2 are given in Table 2.1.

The reason for this particular selection of true parameter values and distributional assumptions is that they guarantee a good mix of the two regimes, i.e. with respect to model A, C and D approximately half of the observations on q_t are determined by the demand equation while the other half is determined by the supply equation. All models are estimated with the maximum likelihood method, for which we used the Marquardt iterative optimisation algorithm. The starting values required for this algorithm are drawn from uniform distributions with the lower and the upper bounds equal to 90% and 110% respectively of the true parameter values.

In order to evaluate the performance of the different estimators we employ the mean absolute deviation (MAD) from the true values. We also accounted for computational problems that occurred during the experiments and which are presumably associated with unbounded likelihood functions and false maxima. Since the estimates in these cases are not reliable we discarded them in order not to distort our results.

2.3.2 Results of Misspecification Experiments

The results of our misspecification analysis can be summarised as follows. In the first group of experiments we find that neither the equilibrium estimator nor estimator A perform well if the data have been generated by the respective other model – although in one instance the equilibrium estimator outperforms estimator A (see Table 2.2). Nevertheless, it seems to be more costly to estimate model A with the equilibrium estimator than to estimate the equilibrium model with estimator A: while the median MAD of estimator A under the true equilibrium model equals 0.20 – for the equilibrium estimator we get 0.08 – the median MAD of the equilibrium estimator under the true model A equals 1.17 while for estimator A we get 0.14.¹² Furthermore, the equilibrium estimator encounters a lot more computational problems if the data have been generated by model A than estimator A encounters with equilibrium data.

Our results for the second group of experiments show that the performance of estimator C under the true model A is in fact highly sensitive to the starting value provided for γ , meaning that the degree of misspecification decreases for lower values (see Table 2.3). Estimator C even outperforms es-

¹²Qualitatively, Goldfeld & Quandt (1981) observe a similar result with regard to estimator C and D and the equilibrium estimator.

estimator A for a very low starting value: in experiment 2.3 estimator C yields lower mean absolute deviations than estimator A in all instances. The starting value of γ also seems to influence the performance of estimator C with regard to computational problems, i.e. less problems occur for lower values. A completely different picture arises in experiment 2.4 to 2.6 in which we focus on the performance of estimator D under the true model A (see Table 2.4). Here the degree of misspecification seems to be independent of the starting value provided for γ .

Let us now turn to the results of the misspecification experiments in which model A is the misspecified model and the data have been generated either by model C or D. From a practical point of view, these results are probably most relevant because a large number of studies in the disequilibrium literature rely on model A. In case the true model is C we observe that the degree of misspecification of A in fact increases with γ although only slightly (see Table 2.5). In terms of lower mean absolute deviations, estimator A never outperforms estimator C. However, for a high starting value of σ_p^2 , which is implicitly equal to zero in model C, estimator A outperforms estimator D. Interestingly, estimator A did encounter hardly any computational problems if the true model is C while estimator D was presumably plagued by the unboundedness problem. Besides, the quality of the estimates provided by estimator C and D decrease as γ decreases. If the data have been generated by model D then the degree of misspecification of model A seems to increase more than under the true model C. In some experiments with a very low γ estimator A even outperforms estimator D (see Table 2.6). Furthermore, for a high true value of σ_p^2 , estimator A outperforms estimator C, which yields very poor estimates in this case. As in all other experiments, estimator A did encounter hardly any computational problems – except for experiment 4.4 in which all estimators were plagued by computational problems. Similar to our observation in the third group, we find that estimators C and D perform poorly for low values of γ .

In summary, we can maintain the following. In case the true model is either model C or D, the degree of misspecification of model A is an increasing function of γ . Nevertheless the estimates provided by estimator A reach a reasonable level of accuracy only for very low true values of γ . Besides, these results are more pronounced in case the data have been generated by model D. If A is the true model and we estimate model C and D respectively, we observe that the performance of estimator C is highly sensitive to the starting value provided for γ , meaning that for very low values estimator

C performs quite well, while estimator D is not. Furthermore, we have the impression that it is more costly to estimate either model C or D under the true model A than to estimate model A under the true model C and D. A similar result seems to be valid for model A and the equilibrium model. Finally, we find that the estimation of model A is associated with surprisingly few computational problems irrespective of the true model.

2.4 Dynamic Models

Another topic that has attracted a lot of attention in the disequilibrium literature is the estimation of disequilibrium models with dynamic features. The reason is that disequilibrium models are usually applied to time series data. This implies that basic assumptions used for the derivation of the likelihood functions of the canonical models, such as no autocorrelation in the error terms and no lagged endogenous explanatory variables, are likely to be violated. However, the estimation of models with unknown sample separation becomes extremely difficult or even impossible if these assumptions are relaxed. For instance, although one can derive the appropriate likelihood function for model A under the assumption of autocorrelated errors, the function is very complex since it involves T -fold integrals with 2^T summations. For model C, in contrast, where sample separation is known, one can derive a likelihood function that is both tractable and accounts for autocorrelation in the disturbances as well as lagged endogenous variables.

Many studies in the disequilibrium literature have therefore concentrated on estimators that account for dynamic features and that are based on tractable likelihood functions. Most of them use various simulation-based estimation methods. Laroque & Salanie (1993), for instance, propose a simulated pseudo maximum likelihood (SPML) method. This approach represents a limited information method which is not based on the entire distribution of the endogenous variables but only on its first two conditional moments. The approach is simulation-based because the two moments are simulated by their empirical counterparts. Lee (1997), in turn, uses a simulated maximum likelihood (SML) approach, in which intractable expectation terms in the likelihood function are replaced with counterparts from Monte Carlo simulations – note that the derived estimator is valid for model A only. According to Fermanian & Salanie (2004), the SPML approach can be applied to many models but it does not yield efficient estimators, while the SML method yields

efficient estimators but can only be applied to a restrictive class of models. Although preliminary results from Monte Carlo experiments show that these estimators yield relatively good estimates, a definitive assessment of these alternative methods is still missing. For instance, both methods described above have merely been tested with data generated by model A, i.e. it is unknown how the SPML method, for instance, performs in a setting where the price is endogenous. Furthermore and due to the fact that some of these simulation-based estimation methods are rather difficult to implement, they are not yet established in the applied disequilibrium literature and so there are no reports regarding experiences with these methods.

Our analysis, on the other hand, is based on a result by *Gourieroux et al. (1985)* who show for a general class of models, including disequilibrium models with unknown sample separation, that the maximum likelihood estimator derived from a static model remains consistent even if the disturbances are serially correlated. As noted by *Quandt (1988)*, however, this result will not hold generally in case of lagged endogenous variables while *Laroque & Salanie (1993)* indicate that consistency should hold for the case of a linear model with normal disturbances. Given these results regarding the asymptotic properties of the static estimators, we analyse by means of sampling experiments the properties of the static estimators A and D in case the data generating process exhibits dynamic features.

2.4.1 Design of Dynamic Model Experiments

In our dynamic model experiments the true models A and D respectively either involve autocorrelated errors or lagged endogenous variables or both – except for experiment 5 and 6 which serve as benchmark. An overview of these experiments is given in Table 2.7 in the appendix. Our artificial data are generated by the following models:

$$\begin{aligned}
d_t &= \alpha_1 + \alpha_2 x_{1,t} + \alpha_3 p_t + \alpha_4 d_{t-1} + \epsilon_{d,t} \\
s_t &= \beta_1 + \beta_2 x_{2,t} + \beta_3 p_t + \beta_4 s_{t-1} + \epsilon_{s,t} \\
q_t &= \min(d_t, s_t) \\
p_t &= p_{t-1} + \gamma(d_t - s_t) + \epsilon_{p,t} \quad (\text{for Model D}) \\
\epsilon_{d,t} &= \rho_1 \epsilon_{d,t-1} + u_{d,t} \\
\epsilon_{s,t} &= \rho_2 \epsilon_{s,t-1} + u_{s,t} \\
\epsilon_{p,t} &= \rho_3 \epsilon_{p,t-1} + u_{p,t} \quad (\text{for Model D})
\end{aligned}$$

where $\alpha_1 = 2$, $\alpha_2 = 1$, $\alpha_3 = -0.5$, $\beta_1 = 1.4$, $\beta_2 = 0.8$, $\beta_3 = 0.1$ and $\gamma = 0.5$. The true parameter values for α_4 and β_4 as well as the autocorrelation coefficients ρ_1 , ρ_2 and ρ_3 are given in Table 2.7. From the table one can see that in experiment 5.1, 5.4, 6.1 and 6.4 the dynamics come from autocorrelated errors while in experiment 5.2, 5.5, 6.2 and 6.5 the errors are serially independent but the explanatory variables include lagged endogenous variables. Finally, in experiment 5.3, 5.6, 6.3 and 6.6 we have both autocorrelated errors and lagged endogenous explanatory variables. The main difference between these experiments is how strongly the dynamic features are pronounced. The disturbance terms $u_{d,t}$, $u_{s,t}$ and $u_{p,t}$, have a normal distribution with mean zero and variance $\sigma_d^2 = 1.44$, $\sigma_s^2 = 1$ and $\sigma_p^2 = 0.72$. In our experiments with autocorrelated errors we choose $u_{d,0} = u_{s,0} = u_{p,0} = 0$. The two exogenous variables x_1 and x_2 are again generated from a uniform distribution over the range (10,15) and (11,17) respectively. As before, for the experiments in which the true model is A, we assume that the exogenous price series is drawn from a uniform distribution over the range (2,5) and that it is fixed for all replications. For experiments with lagged endogenous variables, we also need observations on the initial quantity of demand and supply. We assume that in $t = 0$ the market is in equilibrium and therefore $d_0 = s_0 = 15$. For the initial price level we choose $p_0 = 3.5$. This particular selection of true parameter values and distributional assumptions again guarantees that approximately half of the observations on q_t are determined by the demand equation while the other half is determined by the supply equation. For the estimation of the parameters we once again used the Marquardt iterative optimisation algorithm. The starting values are drawn from uniform distributions with the lower and the upper bounds equal to 90% and

110% respectively of the true values. The properties of the static estimators are then evaluated by means of the mean absolute deviation (MAD) from the true values. As in our misspecification experiments, we set the number of observations to $T = 100$ and the number of replications to $S = 500$ in all experiments. We also account for computational problems as before.

2.4.2 Results of Dynamic Model Experiments

First consider the results regarding model A. Table 2.8 shows that if the dynamics are not too strong then the static estimator performs well even if the errors are autocorrelated or there are lagged endogenous explanatory variables. Most notably, the summary statistics for these two dynamic cases do not differ substantially from those of the static case. If the true model exhibits both dynamic features the performance of the static estimator A gets worse but still seems to be tolerable. For instance, if the true model is static then the median MAD equals 0.15 while in case of autocorrelated errors and lagged endogenous explanatory variables we get 0.23 – for the case of autocorrelated errors and lagged endogenous variables respectively we get 0.17 and 0.16. Furthermore, except for experiment 5.3, we observe that the static estimator does not encounter more computational problems than in the static case.

However, we get a different picture if we look at the experiments in which the dynamics are more pronounced. From Table 2.9 we clearly see that the performance of the static estimator suffers substantially in comparison to the benchmark static case. This fact does not only manifest itself in higher summary statistics but also in an increase in the computational problems encountered during estimation. These results are particularly true for the case where the true model involves lagged endogenous variables. In contrast to the experiments with weak dynamics, the performance of the static estimator in case of autocorrelated errors and lagged endogenous explanatory variables is intolerable, meaning, for instance, that the median MAD equals 0.67 – the corresponding figure for the static case is 0.15. With regard to the discussion surrounding the presence of lagged endogenous explanatory variables we observe that the static estimator A performs better if the error terms are autocorrelated than if the true model involves lagged endogenous variables.

In terms of the dynamic structure experiments for model D we find very similar results as for model A. First, in case of weak dynamics the static

estimator D performs rather well in case of autocorrelated errors or lagged endogenous explanatory variables (see Table 2.10). In addition, the summary statistics for the case of autocorrelated errors and lagged endogenous variables do not differ substantially from the statistics of the static case. As before, however, the more distinct the dynamics of the underlying model get the worse is the performance of the static estimator (see Table 2.11), which becomes apparent not only from an increase in the summary statistics but also from a higher frequency of computational problems encountered during estimation.

In summary, the results from our dynamic structure experiments show that the performances of the static estimators A and D are not convincing if the data have been generated by a model featuring autocorrelated errors and/or lagged endogenous variables. The fact that both estimators perform rather well for less pronounced dynamics is only a small comfort in this respect.

2.5 Conclusion

In this chapter we analyse the consequences and the degree respectively of disequilibrium model misspecification as well as the performance of static disequilibrium estimators in dynamic settings. The sampling experiments conducted for this purpose allow us to draw the following conclusions. While most empirical studies that apply the canonical disequilibrium models have not paid attention to the issue of misspecification, our results suggest that this is a shortcoming. In particular, the consequences of disequilibrium model misspecification are a minor issue only for a very narrow range of the speed of price adjustment. For instance, if the latter is close to zero and the true model is either model C or D then estimating model A can yield reasonable results – primarily if the true model is D. Outside this range, however, misspecification leads to highly biased estimates. Furthermore, our results indicate that researchers should not rely upon the fact that the static estimators for model A and D remain consistent if the data generating model exhibits dynamic features such as autocorrelated errors and/or lagged endogenous variables. If the dynamics are not too pronounced the static estimators actually yield reasonable estimates. However, if the dynamics become stronger this result does not hold anymore. Our findings therefore suggest that researchers should account for model dynamics and different specifications.

2.6 Appendix

Exp.	True Model	Estim. Models	True γ	Starting Value γ	True σ_p^2	Starting Value σ_p^2
1.1	Equ.	Equ./A	—	—	—	—
1.2	A	A/Equ.	—	—	—	—
2.1	A	A/C	—	[0.90,1.10]	—	—
2.2	A	A/C	—	[0.009,0.011]	—	—
2.3	A	A/C	—	[0.00009,0.00011]	—	—
2.4	A	A/D	—	[0.90,1.10]	—	[0.90,1.10]
2.5	A	A/D	—	[0.009,0.011]	—	[0.90,1.10]
2.6	A	A/D	—	[0.00009,0.00011]	—	[0.90,1.10]
3.1	C	C/A/D	1.00	[0.90,1.10]	—	[0.009,0.011]
3.2	C	C/A/D	0.50	[.45,0.55]	—	[0.009,0.011]
3.3	C	C/A/D	0.01	[0.009,0.011]	—	[0.009,0.011]
3.4	C	C/A/D	0.01	[0.009,0.011]	—	[0.90,1.10]
4.1	D	D/A/C	1.00	[0.90,1.10]	0.10	[0.009,0.011]
4.2	D	D/A/C	0.50	[.45,0.55]	0.10	[0.009,0.011]
4.3	D	D/A/C	0.01	[0.009,0.011]	0.10	[0.009,0.011]
4.4	D	D/A/C	0.01	[0.009,0.011]	1.00	[0.90,1.10]

Table 2.1: Misspecification experiments

Experiment	1.1		1.2	
True Model	Equ.		A	
Estimator	Equ.	A	Equ.	A
MAD estim. \leq MAD appropriate estim. (# of instances)	8	0	8	1
MAD estim. / MAD appropriate estim. (median)	1.00	2.34	1.00	2.16
Median MAD	0.08	0.20	0.14	1.17
Average number of iterations	56	42	41	48
Repl. with computational problems	0	13	3	67
Repl. with number of iterations > 500	8	1	1	0

Table 2.2: Results of Misspecification Experiments (Group 1)

Experiment	2.1		2.2		2.3	
True Model	A		A		A	
Estimator	A		C		A	
MAD estim. \leq MAD appropriate estim. (# of instances)	8		1		8	
MAD estim. / MAD appropriate estim. (median)	1.00		15,556.84		1.00	
Median MAD	0.16		3,124.67		1.00	
					0.14	
					1.19	
					0.14	
					0.05	
Average number of iterations	42		56		42	
Repl. with computational problems	4		273		8	
Repl. with number of iterations > 500	1		3		2	
					10	
					2	
					0	

Table 2.3: Results of Misspecification Experiments (2.1 to 2.3)

Experiment	2.4		2.5		2.6	
True Model	A		A		A	
Estimator	A		D		A	
MAD estim. \leq MAD appropriate estim. (nr. of instances)	8		0		8	
MAD estim. / MAD appropriate estim. (median)	1.00		435.34		1.00	
Median MAD	0.14		66.04		1.00	
					856.46	
					1.00	
					1561.30	
					0.13	
					219.03	
Average number of iterations	47		51		43	
Repl. with computational problems	2		53		2	
Repl. with number of iterations > 500	3		0		5	
					0	
					2	
					0	

Table 2.4: Results of Misspecification Experiments (2.4 to 2.6)

Experiment	3.1				3.2				3.3				3.4			
True Model	C				C				C				C			
Estimator	A				A				A				A			
MAD estim. \leq MAD appropriate estim. (nr. of instances)	8	0	3	8	0	3	8	0	3	8	0	3	8	0	3	8
MAD estim. / MAD appropriate estim. (median)	1.00	2.27	1.06	1.00	2.32	1.03	1.00	2.32	1.03	1.00	1.73	1.03	1.00	1.79	2.05	2.05
Median MAD	0.08	0.20	0.09	0.07	0.17	0.08	0.62	1.02	0.63	0.57	0.96	1.34	0.57	0.96	1.34	1.34
Average number of iterations	56	44	40	60	46	36	63	52	43	67	51	46	67	51	46	46
Repl. with computational problems	0	2	259	0	2	270	0	1	227	0	0	152	0	0	152	152
Repl. with number of iterations > 500	8	7	1	12	4	2	5	2	1	17	4	1	17	4	1	1

Table 2.5: Results of Misspecification Experiments (Group 3)

Experiment		4.1			4.2			4.3			4.4		
True Model		D			D			D			D		
Estimator		D	A	C	D	A	C	D	A	C	D	A	C
MAD estim. \leq MAD appropriate estim. (nr. of instances)		8	0	5	8	0	4	8	3	0	8	6	0
MAD estim. / MAD appropriate estim. (median)		1.00	2.46	0.98	1.00	2.24	0.97	1.00	1.02	2.36	1.00	0.97	37,408.70
Median MAD		0.08	0.20	0.07	0.08	0.19	0.08	0.34	0.33	0.75	0.14	0.14	8,886.80
Average number of iterations		36	50	58	36	44	56	45	49	90	40	43	199
Repl. with computational problems		241	1	0	191	3	0	1	51	43	45	43	253
Repl. with number of iterations > 500		3	3	16	0	5	8	1	7	0	5	3	189

Table 2.6: Results of Misspecification Experiments (Group 4)

Exp.	True/Estim.	Model	ρ_1	ρ_2	ρ_3	$\alpha_4 = \beta_4$
5	A		0.00	0.00	—	0.00
5.1	A		0.40	0.50	—	0.00
5.2	A		0.00	0.00	—	0.10
5.3	A		0.40	0.50	—	0.10
5.4	A		0.70	0.80	—	0.00
5.5	A		0.00	0.00	—	0.30
5.6	A		0.70	0.80	—	0.30
6	D		0.00	0.00	0.00	0.00
6.1	D		0.40	0.50	0.30	0.00
6.2	D		0.00	0.00	0.00	0.10
6.3	D		0.40	0.50	0.30	0.10
6.4	D		0.70	0.80	0.60	0.00
6.5	D		0.00	0.00	0.00	0.30
6.6	D		0.70	0.80	0.60	0.30

Table 2.7: Dynamic structure experiments

Experiment	5	5.1	5.2	5.3
Mean MAD	0.49	0.56	0.54	0.76
Median MAD	0.15	0.17	0.16	0.23
Mean MSE	1.42	1.95	1.60	3.42
Median MSE	0.04	0.05	0.04	0.12
Average number of iterations	45	50	50	47
Repl. with computational problems	0	0	0	9
Repl. with number of iterations > 500	7	4	4	3

Table 2.8: Results of Dynamic Structure Experiments (5.1 to 5.3)

Experiment	5	5.4	5.5	5.6
Mean MAD	0.49	0.81	3.83	2.00
Median MAD	0.15	0.31	1.62	0.67
Mean MSE	1.42	3.74	3.314.45	30.18
Median MSE	0.04	0.18	23.36	1.13
Average number of iterations	45	45	50	48
Repl. with computational problems	0	5	251	186
Repl. with number of iterations > 500	7	2	2	0

Table 2.9: Results of Dynamic Structure Experiments (5.4 to 5.6)

Experiment	6	6.1	6.2	6.3
Mean MAD	0.23	0.26	0.34	0.37
Median MAD	0.09	0.10	0.08	0.10
Mean MSE	0.28	0.36	0.70	0.78
Median MSE	0.01	0.01	0.01	0.02
Average number of iterations	44	39	46	42
Repl. with computational problems	0	0	0	0
Repl. with number of iterations > 500	5	5	5	6

Table 2.10: Results of Dynamic Structure Experiments (6.1 to 6.3)

Experiment	6	6.4	6.5	6.6
Mean MAD	0.23	0.41	0.94	1.12
Median MAD	0.09	0.15	0.11	0.21
Mean MSE	0.28	0.82	4.80	14.06
Median MSE	0.01	0.03	0.02	0.07
Average number of iterations	44	38	48	49
Repl. with computational problems	0	0	76	98
Repl. with number of iterations > 500	5	2	2	1

Table 2.11: Results of Dynamic Structure Experiments (6.4 to 6.6)

Chapter 3

Non-Price Credit Rationing by U.S. Commercial Banks During the Great Recession

3.1 Introduction

The bursting of a bubble in the U.S. housing market in early 2007 triggered a global financial and economic crisis not witnessed since the Great Depression.¹ While at first only the financial sector was affected, the crisis quickly spread out into the real economy. Bank lending has been assumed to be an important transmission channel. According to conventional wisdom, the severe disruptions within the financial system led banks to cut back on lending sharply. This rationing of credit, in turn, resulted in a contraction in consumption and investment and economic activity as a whole. For this reason, the term “credit crunch”, which is meant to describe a sharp reduction in bank credit supply, became one of the many names for the crisis.

Although this view about the transmission of the crisis enjoys widespread acceptance it is also subject to controversy. Chari et al. (2008), for instance, provide empirical evidence for the U.S. showing that aggregate lending to nonfinancial borrowers actually increased during the first phase of the crisis.²

¹For a preliminary analysis of the crisis, which is referred to as the Great Recession, see, for instance, Blanchard (2008), Brunnermeier (2009) or Krishnamurthy (2010).

²The results by Chari et al. (2008), however, are based on data up to October 25, 2008, i.e. they only partially reflect the financial turmoil that followed the collapse of the investment bank Lehman Brothers.

The findings of Ivashina & Scharfstein (2010), on the other hand, who analyse the U.S. market for syndicated loans are consistent with the credit crunch hypothesis. They admit, however, that the observed contraction in lending could also be explained with a reduced demand for credit due to cyclical weakness – according to the National Bureau of Economic Research, the recession in the U.S. already started in the fourth quarter of 2007. Contessi & Francis (2009), in turn, show that credit flows in the fourth quarter of 2008 resembled the behaviour during the 1990-91 recession which is commonly associated with a credit crunch.

Against this background this chapter investigates if the U.S. market for commercial and industrial (C&I) loans exhibits signs of a supply-induced credit crunch. In particular, we analyse whether U.S. commercial banks increased non-price credit rationing sharply during the financial crisis that started in 2007. Our analysis is based on a disequilibrium econometric framework that enables us to quantify the extent of (the change in) credit rationing in the U.S. market for C&I loans. The main innovation in this respect is the application of an excess demand indicator other than the change in the loan rate with which we can relate observations on the credit quantity to a demand and a supply regime respectively. In contrast to previous studies, this classification not only allows us to estimate a rich array of dynamic disequilibrium models but also to unify two hitherto different approaches in the empirical credit rationing literature.

The remainder of this chapter is structured as follows. Section 3.2 provides a short review of the empirical literature on credit rationing with a particular emphasis on the U.S. credit market. It also deals with the associated literature on credit crunches. In section 3.3 we present our empirical models and we undertake a preliminary analysis of the applied excess demand indicators. The section closes with a presentation of our estimation results. Finally, section 3.4 concludes.

3.2 Literature Review

In general, one has to distinguish between (conventional) price and non-price rationing. The former refers to the allocation of scarce goods and services using prices, i.e. only agents who are willing and able to pay the market price can purchase a particular good or service while all other agents are rationed. In the following, however, we are mainly concerned with non-price rationing.

For simplicity, the latter is referred to as rationing throughout the text.

With regard to credit, rationing is usually defined as a situation in which the demand for credit exceeds the supply at the ruling market price, i.e. the demand of some borrowers is unfulfilled although they would be willing to pay the market price. As noted by Baltensperger (1978), the meaning of the term price is a subtle issue in this respect. Under the narrow, more traditional definition of rationing the market price represents a vector that includes the interest rate as well as other loan terms, such as collateral requirements. The broader definition, on the contrary, equates the market price with the interest rate.

Furthermore, one has to distinguish between “equilibrium” and “disequilibrium” rationing. The latter is a temporary phenomenon that occurs during the adjustment process from one (Walrasian-) equilibrium to another. It is associated with situations after an exogenous shock in which the price, i.e. the loan rate, does not instantaneously adjust to balance credit demand and supply due to price stickiness. Equilibrium rationing, in contrast, is defined as a permanent phenomenon that is due to information asymmetries in the market.³ It describes

“[...] circumstances in which either (a) among loan applicants who appear to be identical some receive a loan and others do not, and the rejected applicants would not receive a loan even if they offered to pay a higher interest rate; or (b) there are identifiable groups of individuals in the population who, with a given supply of credit, are unable to obtain loans at any interest rate, even though with a larger supply of credit they would.” (Stiglitz & Weiss 1981)

While the contributions to the theoretical credit rationing literature are numerous, empirical evidence regarding this phenomenon is relatively scarce. The main problem is that credit demand and supply – and thus the difference between the two – are not directly observable.⁴ For this reason various indirect approaches – based on proxy measures, survey data or testable implications – have been developed. Although these approaches are capable of

³In our view, though, the boundary between equilibrium and disequilibrium rationing is blurred. A case which supports our view is where the price adjusts only marginally to imbalances between demand and supply.

⁴Available data include the actual credit quantity transacted in the market – or more precisely the quantity of loans outstanding.

providing evidence for the existence of credit rationing they cannot quantify its extent. The only direct approach with which one can estimate excess demand and excess supply respectively are so-called disequilibrium econometric models.

3.2.1 Indirect Approaches to Credit Rationing

Jaffee & Modigliani (1969) were among the first to apply the proxy method. They argue that the (unobservable) degree of rationing, defined as the ratio of rationed credit to potential demand from rationed customers, is closely related to an (observable) proxy measure given by the ratio of loans granted to risk-free customers – approximated by loans granted at the prime rate – to total loans. In particular, they show that the proxy measure is a monotonic function of the degree of rationing. Using this proxy, they provide evidence that there exist considerable variations in the degree of rationing over time in the U.S. market for C&I loans.⁵

A typical exponent of the survey method, which plays an important role in our own analysis, is Harris (1974). He argues that credit rationing is based on the allocation of credit via non-price criteria. As a result, rationing is closely associated with changes in non-price loan terms, i.e. changes in credit standards of banks. Positive rationing (or excess demand) occurs when banks increase non-price requirements on loans while an easing of requirements implies negative rationing (or excess supply).⁶ For his analysis of the U.S. market for C&I loans Harris uses data from the Senior Loan Officer Opinion Survey on Bank Lending Practices conducted every quarter by U.S. monetary authorities. In this survey, loan officers from a representative selection of U.S. commercial banks are asked whether their credit standards have changed over the past three months and how they would characterise this change – basically, the categories are tightened, unchanged and eased. The net percentage of respondents reporting a tightening of standards, i.e. the difference between the percentage of those reporting a tightening and those reporting an easing of standards, then serves as an indicator of credit rationing. While a positive balance indicates that credit standards have been

⁵One problem with this approach is that the proxy does not only vary with the degree of rationing but also because of changes in another parameter, which is not necessarily constant over time.

⁶Note that according to the classification of Baltensperger (1978) this definition of rationing falls into the broad category.

made more stringent across banks implying positive credit rationing, a negative balance indicates negative rationing. Using survey data from 1964 to 1970, Harris is able to provide evidence for the existence and variability of credit rationing in the U.S.

Berger & Udell (1992), in turn, use testable implications to find empirical evidence for rationing. Although their results suggest that the commercial loan rate in the U.S. is sticky as with disequilibrium rationing, they argue that this stickiness reflects rationing only partially. Their argument is based on two main findings: first, half of the loan rate stickiness occurs on loans made under a commitment. These loans, however, are contractually protected from rationing. And second, the ratio of loans made under commitment and loans made under no commitment should rise in times of tight credit, i.e. in times when credit is rationed. Their results suggest that this is not the case. Therefore they conclude that although credit rationing may exist it probably does not constitute an important macroeconomic phenomenon.

3.2.2 Disequilibrium Econometric Models

As already indicated above, unlike the indirect approaches disequilibrium models can be used to estimate excess demand directly and thus to quantify the extent of credit rationing. The most basic specification of a disequilibrium model is given by the following system of equations:

$$d_t = \alpha_1 p_t + \boldsymbol{\alpha}_2' \mathbf{x}_{d,t} + \epsilon_{d,t} \quad (3.1)$$

$$s_t = \beta_1 p_t + \boldsymbol{\beta}_2' \mathbf{x}_{s,t} + \epsilon_{s,t} \quad (3.2)$$

$$q_t = \min(d_t, s_t) \quad (3.3)$$

where d_t denotes demand in period t , s_t is supply, q_t is the actual quantity transacted in the market, p_t is the market price associated with the parameters α_1 and β_1 respectively, $\mathbf{x}_{d,t}$ and $\mathbf{x}_{s,t}$ are vectors of exogenous variables, $\boldsymbol{\alpha}_2$ and $\boldsymbol{\beta}_2$ are parameter vectors and $\epsilon_{d,t}$ and $\epsilon_{s,t}$ are disturbance terms. The latter are usually assumed to be jointly normally distributed as well as serially and mutually independent.

The most distinguishing feature of this model is that instead of an equilibrium condition of the form $q_t = d_t = s_t$ it includes a minimum condition implying that the actual quantity transacted in the market is the minimum

of demand and supply. Hence, it is not assumed that the price clears the market. The minimum condition (3.3) is usually justified on the grounds that most markets are characterised by voluntary exchange, i.e. neither side of the market is forced to trade more than it wishes. Since the price is assumed to be exogenous in equations (3.1) to (3.3), i.e. the price is fully rigid in the sense that it does not adjust to imbalances between demand and supply, the system is often augmented by a price adjustment equation to make the price explicitly endogenous. The adjustment equation usually takes one of the following two forms:

$$p_t = p_{t-1} + \gamma(d_t - s_t) \quad (3.4a)$$

$$p_t = p_{t-1} + \gamma(d_t - s_t) + \epsilon_{p,t} \quad (3.4b)$$

Although price adjustment is deterministic in equation (3.4a) and stochastic in equation (3.4b), the price in period t is in both cases endogenous and a function of the quantity of excess demand, i.e. it adjusts to market imbalances, though it does not necessarily offset them. The parameter γ captures the speed of price adjustment and is assumed to be positive. In case of the stochastic price adjustment equation the disturbance terms of the system are commonly assumed to be joint normal as well as serially and mutually independent.

Equations (3.1) to (3.3) are usually referred to as model A while equations (3.1) to (3.4a) and (3.1) to (3.4b) are referred to as model C and model D respectively. The three models, representing the canonical disequilibrium models, differ from each other in some important respects. First, in model A and D sample separation is unknown while in model C it is known.⁷ Sample separation means that the minimum condition separates the observations for q_t into observations which are determined by the demand equation and observations which are determined by the supply equation. A lack of sample separation information has a number of important implications for estimation which we discuss in more detail in section 3.3.1. Furthermore, whether the price is assumed to be exogenous or endogenous is closely related to the type of credit rationing that is associated with each model. Model C and D are based on the assumption of price stickiness, i.e. although in the short-run the price maybe different from the equilibrium level implying rationing, in

⁷Note that in model C the sign of $\Delta p_t = p_t - p_{t-1}$ and thus the sign of $(d_t - s_t)$ and the minimum are known.

the long-run it is assumed to return to its equilibrium level. Hence, one can analyse disequilibrium or temporary rationing with these two models. Model A, on the other hand, captures the essence of equilibrium rationing since the price is not assumed to change due to imbalances between supply and demand.

The disequilibrium method was first introduced in a seminal paper by Fair & Jaffee (1972). Since then a number of studies have applied disequilibrium models to credit markets. Sealey (1979), for instance, applied model D to U.S. commercial & industrial loan market data. Analysing the time period from 1952 to 1977, he finds that the market for C&I loans is characterised by disequilibrium rationing and that the magnitude of the imbalance between credit supply and demand can be quite large. His results, which are consistent with the findings of Harris (1974), further indicate that the speed of price adjustment is rather low. Furthermore, they suggest that the U.S. credit market is characterised by both periods of excess demand and periods of excess supply.

In another study, Ito & Ueda (1981) use a variant of model C proposed by Bowden (1978) to analyse the adjustment speed of the loan rate in the U.S. and the Japanese credit market respectively.⁸ Their results indicate that in the U.S. the prime loan rate adjusts rapidly to imbalances between credit demand and supply and that there is no significant difference between the upward and the downward adjustment speed, i.e. the interest rate does not react differently to excess demand and excess supply. The former result implies that one cannot reject the hypothesis that demand is always equal to supply in the U.S. credit market.

King (1986), meanwhile, focuses on equilibrium rationing in the U.S. market for C&I loans. Applying model A, he finds mixed support for the hypothesis of equilibrium credit rationing. In contrast to what the equilibrium rationing model would predict, his results suggest that the loan rate elasticity of supply is significantly different from zero. However, he interprets the fact that the C&I loan market seems to be predominantly in a state of excess demand as evidence for equilibrium rationing.

The study by Kugler (1987), in which the credit markets in Switzerland, West Germany, the United Kingdom and the U.S. are analysed using the

⁸In the specification suggested by Bowden (1978), the price in period t is a weighted average of the price in the previous period and the equilibrium price in period t . According to some authors, this specification offers a number of advantages in terms of testing the disequilibrium hypothesis against the equilibrium hypothesis.

Bowden (1978) specification, again concentrates on the speed of price adjustment. Based on his results, Kugler cannot reject the equilibrium hypothesis for the U.S., which is consistent with the findings of Ito & Ueda (1981). Furthermore the hypothesis of equal upward and downward adjustment speeds cannot be rejected in all four countries.

Finally, in his empirical analysis of credit rationing in the U.S. commercial & industrial loan market, Mayer (1989) argues that the notion that the price tends to change in the direction of excess demand – incorporated in deterministic price adjustment equations such as (3.4a) – is too vague to be used in the estimation of credit demand and supply. Although he indicates that stochastic adjustment equations such as (3.4b) remedy this deficiency, he claims that models including the latter are computationally much more difficult to handle and therefore suggests to use a different model. The price adjustment process he proposes is based on the notion that the probability of a price increase is higher in case of excess demand than in case of excess supply. For this approach, it is not necessary to specify explicitly a price adjustment equation and the likelihood function of the new model is computationally simpler.⁹ Using data from 1979 to 1984 he estimates credit demand and supply showing that there is a reasonable pattern of excess demand. His results are consistent with those of Sealey (1979).

3.2.3 Credit Rationing and Credit Crunches

In addition to the papers presented in the previous section that concentrate on credit rationing per se, in recent years a number of studies have used the disequilibrium econometric framework to analyse so-called credit crunch episodes. As we point out below, these episodes can involve non-price credit rationing. Unfortunately, though, the academic literature differs about this connection. In general,

“The term ‘credit crunch’ has come to be applied indiscriminately [...] to describe any and all conditions of expensive or difficult-to-obtain credit. Such usage provides a catchy label, but one with no particular content. (Wojnilower 1994)

The controversy about the term credit crunch mainly refers to the manifestations of this phenomenon. In a seminal paper on the subject Bernanke

⁹Mayer (1989) admits, however, that some rather restrictive assumptions are necessary to estimate this new model.

& Lown (1991) define a crunch as a significant leftward shift in the supply curve for bank loans that causes a sharp increase in price rationing and which does not necessarily involve quantity or non-price rationing. Other scholars like Friedman (1991), however, doubt that a simple leftward shift or a mere increase in price rationing would qualify as a crunch. A number of authors have tried to reconcile these two views by defining a credit crunch as a reduced willingness to lend which manifests itself either in the form of a sharp increase in price or non-price rationing.¹⁰ We adhere to this latter, more comprehensive definition of a crunch.

The above described disagreement has produced different empirical approaches to analyse potential credit crunch episodes. In the following we focus on studies that relate a crunch to non-price credit rationing.¹¹ Examples in this respect are Pazarbasioglu (1997), Gosh & Gosh (1999), Kim (1999), Barajas & Steiner (2002), Ikhida (2003), Nehls & Schmidt (2003), Allain & Oulidi (2009) and Poghosyan (2010). These studies apply the disequilibrium framework – most often model A – to loan market data from various countries, excluding the U.S. In large part their results indicate that the loan markets under investigation are characterised by credit rationing. With the exception of Nehls & Schmidt (2003), most of these studies regard non-price credit rationing per se as an indicator for a credit crunch. The definition of a crunch, however, suggests that non-price rationing is not a sufficient condition for a credit crunch. Theoretical as well as empirical evidence indicate that non-price rationing constitutes a usual phenomenon in credit markets which is consistent with normal bank behaviour and that the degree of rationing is likely to change over time and the business cycle. Thus, one has to identify time periods in which excess demand for credit was exceptionally high in order to provide evidence for a crunch. Therefore the aim of the subsequent empirical analysis is to analyse whether U.S. commercial banks increased non-price credit rationing sharply during the global financial crisis that started in 2007.

¹⁰See, for instance, Akhtar (1993).

¹¹Bernanke & Lown (1991), Berger & Udell (1994), Peek & Rosengren (1995) and Brinkmann & Horvitz (1995) are examples of studies that analyse credit crunch episodes from a different perspective.

3.3 Empirical Analysis

3.3.1 Preliminary Discussion

In section 3.2.2 we have pointed out that a lack of sample separation information has important implications for the estimation of disequilibrium models. For instance, the estimation of model A and D respectively, in which sample separation is unknown, is restricted to a few methods of which maximum likelihood is the most widely-used.¹² For the latter one usually has to provide sensible starting values, which is a difficult task. In addition, it is well known that the likelihood functions of models with unknown sample separation are ill-behaved. Maddala (1983), for instance, shows that for certain parameter values the functions are unbounded rendering maximisation impossible. According to Quandt (1988), this is the result of the substantial latency in these models: not only are the quantities demanded and supplied unobserved but sample separation is also unknown.

Accounting for dynamics in models with unknown sample separation poses another, even more severe problem. The issue is that after loosening the basic assumptions of the canonical models, that is no autocorrelated disturbances and no lagged endogenous variables, estimation becomes extremely difficult since the resulting likelihood functions are very complex – an example would be the likelihood function for model A under the assumption of autocorrelated errors. However, estimating dynamic disequilibrium models is important for several reasons, including the fact that they are usually applied to time series data which is why the basic assumptions mentioned above are likely to be violated.

In addition, Browne (1987) notes that most disequilibrium studies that analyse credit markets implicitly assume that there are no adjustment costs to the quantity of credit. While this might be true for credit flows, it is hard to argue that banks as well as borrowers can adjust their stock of outstanding credit instantaneously and without cost. Since data on credit flows are not (publicly) available and one has to work with stock variables, adjustment costs should be accounted for by including lagged credit quantities as explanatory variables in both the demand and the supply equation as suggested by Ito & Ueda (1981). According to Browne (1987), the exclusion of the lagged quantities potentially leads to wrong conclusions about the extent

¹²More information on alternative estimation methods, like non-linear least squares, is provided in Quandt (1988) and Srivastava & Rao (1990).

of rationing. This is because short-run imbalances between credit demand and supply could erroneously be attributed to dynamic credit rationing although they are a consequence of quantity adjustment costs.

While Gouriéroux et al. (1985) have shown that the estimators of static disequilibrium models with unknown sample separation remain consistent even if the disturbance terms are autocorrelated, Quandt (1988) notes that this result not necessarily holds for the case of lagged endogenous variables. Recent contributions in the disequilibrium literature have thus tried to develop estimation methods which are both tractable and which account for various kinds of model dynamics. Examples in this respect are Laroque & Salanie (1993) and Lee (1997) who propose simulation based maximum likelihood estimation methods.

We, in turn, adopt a different strategy which is based on the fact that models with known sample separation do not share the problems outlined above. Apart from the maximum likelihood method, model C, for instance, can be estimated with the standard two-stage least squares procedure. Furthermore it is possible to estimate dynamic versions of the model with autocorrelated errors and/or lagged endogenous variables. However, some authors like Mayer (1989) argue that the deterministic price adjustment equation in model C is too strong an assumption and that the change in the price is probably not a good excess demand indicator.

In terms of the U.S. market for C&I loans and the associated loan rate this criticism seems to be justified for the following reason. The deterministic price adjustment equation (3.4a) not only implies that excess demand is proportional to the change in the price but that excess demand is the major determinant of the price. Figure 3.1, however, suggests that this is not the case for the loan rate. Instead the bank prime loan rate (BPLR) closely follows the U.S. federal funds effective rate (FFER). Hence, changes in the loan rate are not due to excess credit demand but to changes in the interest rate target of the U.S. Fed.

In order to account for the criticism surrounding price adjustment equations, an indicator other than the change in the loan rate is needed. The survey method by Harris (1974) provides such an indicator.¹³ Although it remains difficult to determine whether the regime classification information provided by this alternative indicator is correct and hence the estimates are

¹³The indicator proposed by Harris (1974) has already been used successfully in the past by Lown & Morgan (2006), though in a somewhat different context.

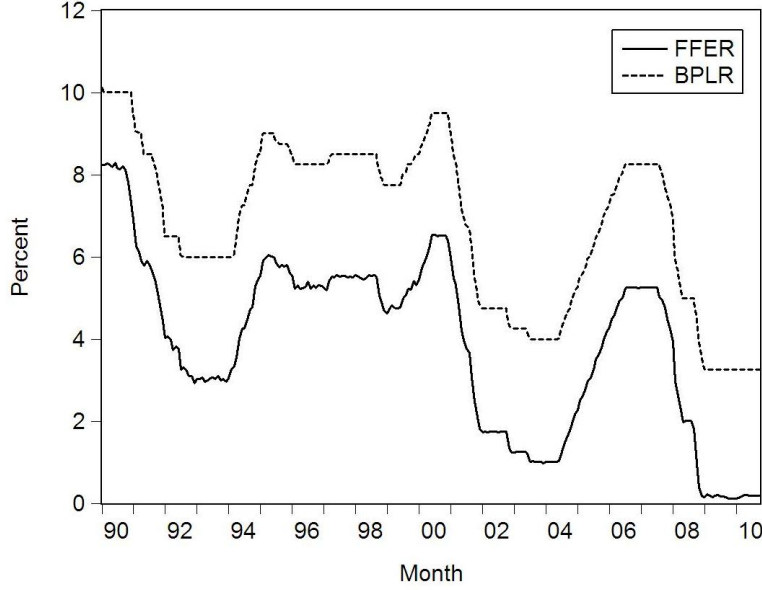


Figure 3.1: Loan Rate and Federal Funds Rate

sufficiently reliable, Rudebusch (1987) notes that the potential bias resulting from a non-perfect regime classification should be considered together with the large increase in model tractability. In fact, in combining the survey method with the disequilibrium approach it is possible to estimate models with a priori unknown sample separation with a wide array of methods and without worrying about unbounded likelihood functions and other difficulties while at the same time model dynamics can be taken into account. Thus, in contrast to models without any sample separation information one can confront many sources of misspecification that would bias the estimation results.

3.3.2 Models with Sample Separation Information

In our analysis of credit rationing in the U.S. market for C&I loans we proceed as follows. First we estimate a dynamic version of model C. In a second step we estimate models with the excess demand indicator proposed by Harris (1974). In particular equations (3.1) to (3.3) are augmented by the following indicator equation

$$d_t - s_t = \delta(z_{t+1} - z^*) \quad (3.5)$$

where d_t and s_t are credit demand and supply respectively, the indicator variable z_{t+1} is the net percentage of banks reporting a tightening of credit standards in the Senior Loan Officer Opinion Survey and z^* is the equilibrium value of the indicator – the subsequent discussion makes it clear that $z^* = 0$. Since the survey in period $t+1$ refers to changes in credit standards in period t , the appropriate time index of the indicator in equation (3.5) is $t+1$. The parameter δ , which is assumed to be positive, is a factor of proportionality, i.e. excess demand for credit is proportional to the net percentage of banks tightening credit standards.

If the net percentage of banks reporting a tightening of standards is positive, i.e. if $z_{t+1} > 0$, then credit demand in period t is bigger than credit supply, i.e. $d_t > s_t$, meaning that there is positive credit rationing. The underlying intuition is the following. If the percentage of banks tightening credit standards for C&I loans gets bigger than the percentage easing them, a certain fraction of firms, who would have got a business loan before, are now quantity rationed although they would be willing to pay the market rate on C&I loans. In case that $z_{t+1} < 0$ the argument is similar in nature.

In terms of estimation, the main advantage of equation (3.5) becomes evident if it is combined with the minimum condition (3.3). For instance, if $s_t < d_t$ one can infer that $s_t = q_t$ and $d_t = q_t + \delta z_{t+1}$. Thus, a system of simultaneous equations with unobserved endogenous variables, i.e. d_t and s_t , is easily transformed into a system where all variables are observed – lagged endogenous credit quantities are handled in the same fashion.¹⁴ Disequilibrium models transformed in this way can be estimated by standard least squares procedures for simultaneous equation models.¹⁵ Also note that in the augmented model the price is a priori exogenous. Since this is a rather restrictive assumption we estimate two variants of the model: one in which the price is exogenous and one in which it is implicitly endogenous. This approach has the advantage that we do not have to specify a price adjustment equation explicitly in order for the price to be endogenous.

¹⁴Note that in model C the transformation works in a similar way. This can be seen from equation (3.4a). Given that $s_t < d_t$, one can infer that $s_t = q_t$ and $d_t = q_t + (1/\gamma)\Delta p_t$.

¹⁵For more details on the transformed systems see the appendix.

3.3.3 Demand & Supply Specification

In order to complete the models outlined in the previous section we have to specify the demand and the supply equation respectively. In doing so we adhere to previous studies presented in section 3.2.2.

With regard to the factors determining credit supply we note the following.¹⁶ Given the amount of loanable funds, in particular deposits, banks have to decide how to allocate these funds. The portfolio decision whether to originate loans or to invest in other assets mainly depends on risk-return considerations. If the spread between the loan rate and the rate on alternative assets, such as government securities, widens it becomes, *ceteris paribus*, more profitable to invest in loans. The decision to originate loans is further strengthened by a favourable economic outlook, which makes it less likely that borrowers cannot repay their loans. Finally, one should account for quantity adjustment costs and the costs of loanable funds. Thus, bank loan supply is specified in the following way:

$$s_t = g(r_{L,t} - r_{T,t}, TD_t, IP_{t-1}, CD_t, s_{t-1}) \quad (3.6)$$

where r_L is the rate on short-term business loans, r_T is the rate on alternative assets represented by government treasury bills, TD is the amount of total deposits, IP is an index of industrial production reflecting expectations about the economic outlook, CD measures the costs per dollar of deposits and s_{t-1} accounts for quantity adjustment costs and other, unobserved factors. The reason, why we use the lagged value of the index of industrial production is that one should not use a proxy for future economic activity that potentially reflects credit rationing, like investment in period t . Prior expectations about these factors determining credit supply indicate that except for the effect of the variable CD , which is indeterminate, all factors should affect supply positively.¹⁷

In terms of the demand specification, three decisions made by firms are considered. They concern the questions whether firms should invest at all, whether firms should finance their investments with internal or external funds

¹⁶In terms of the supply specification, many studies refer to Melitz & Pardue (1973).

¹⁷On the one hand we would expect that an increase in the costs per dollar of deposits reduces banking activity as a whole, implying less credit supply. However, an increase in the variable CD also indicates that banks are less dependent on demand deposits and that their operations are mainly funded by time deposits. Hence the liquidity needs by banks decrease and therefore they can invest their funds in more illiquid assets such as loans.

– note that firms also need funds to finance their ongoing operations, i.e. working capital – and finally which type of external financing they should use. Unfortunately, there is no theoretical framework for the analysis of this interrelated decision making process of firms on which we could base our specification – the supply specification, in contrast, is largely based on the microeconomic theory of the banking firm. Nevertheless, we can maintain the following. Firms are likely to invest and to produce more respectively if their expectations about future economic activity are positive. Furthermore, if firms produce at full capacity, they are likely to invest in new equipment. Whether firms finance their investments and operations internally or externally depends on the amount (and cost) of internal funds. And finally, if firms decide to use external funds, the demand for credit is affected by the cost of alternative forms of external financing, e.g. bonds or commercial paper.¹⁸ Therefore we specify demand as follows:

$$d_t = f(r_{A,t} - r_{L,t}, PT_{t-1}, IP_{t-1}, CU_{t-1}, d_{t-1}) \quad (3.7)$$

where r_L is again the interest rate on short-term business loans, r_A is the interest paid on alternative forms of external financing, PT are corporate profits after taxes, which are a proxy for internal funds, IP is the index of industrial production and CU is capacity utilisation. We also include d_{t-1} to account for quantity adjustment costs and other, unobserved factors. The variables PT , IP and CU are lagged by one period to allow for a time lag in the decision making process of firms. A priori we would expect that all variables, except for profits, affect demand positively.¹⁹

3.3.4 Data & Preliminary Analysis

For our estimations we use data from the U.S. Federal Reserve System.²⁰ For a detailed description of the different variables see Table 3.4 in the appendix.

¹⁸Notice that for small and to a certain extent also medium sized firms, bank loans still represent the sole form of external financing, while for large firms there exist various alternatives.

¹⁹The effect of profits after taxes is ambiguous for the following reason. On the one hand, if firms have higher profits they are less dependent on bank credit. On the other hand, banks are more likely to lend to firms if they have higher profits.

²⁰The data we use for our analysis can be obtained via the Data Download Programme on the website of the Board of Governors of the Federal Reserve System (<http://www.federalreserve.gov/econresdata/releases/statisticsdata.htm>).

We use quarterly data from 1990:q1 to 2010:q2 – in case that only monthly data are available we calculate three-month averages. Our choice of this particular sample period is mainly governed by the fact that there is a break in the indicator series z since banks have not been asked about the tightening of credit standards from 1984 to 1990.

Except for the interest rates and the indicator z all data are seasonally adjusted. The nominal variables q , i.e. the quantity of loans outstanding, TD and PT are normalised by the GDP deflator with base year 2005.²¹ Interest rates are measured in percentage points and have been adjusted for inflation by subtracting the annual increase in the GDP deflator. The reason for using real data is simply that we do not want our results to be distorted by changes in the price level.

Also note that the credit quantity q used throughout our analysis is a stock variable measuring the nominal amount of loans outstanding. We use this variable because there are no publicly available data on credit flows – data on individual C&I loans from the Survey of Terms of Business Lending of the FED, for instance, are not only classified information but they are also gathered infrequently. For this reason the credit demand and supply specifications from above should be interpreted as describing the desired amount of bank debt firms want to hold and banks are willing to offer.

The variable CD , representing the costs per dollar of deposits, is calculated as follows. First we multiply the value of time and savings deposits with the 6-month (real) interest rate on certificates of deposits. The latter is a proxy for the rate paid on time and savings deposits. This gives the total costs for time and savings deposits. These costs are then divided by the value of total deposits to get the costs per dollar of deposits. The CD variable was suggested by Melitz & Pardue (1973). They argue that the costs for checkable and demand deposits respectively are usually offset by service fees charged by banks and therefore the costs for deposits only accrue from time and savings deposits.

It is also worth mentioning that the indicator variable z measures the net percentage of banks tightening credit standards for large and medium sized firms. Although there exists another series referring to the tightening of credit standards for small firms, it is irrelevant which series one uses because they move closely together and have a correlation coefficient higher than 0.95.

²¹Nominal and real GDP data are from the website of the U.S. Bureau of Economic Analysis (<http://www.bea.gov/national/index.htm>).

A much more serious issue is the fact that the time series used for our analysis are integrated of order one.²² This implies that estimations using levels could face spurious regression problems. While disequilibrium studies in the past did not account for this issue, we estimate our models in first differences.

Also note that our results regarding rationing are mainly driven by the applied excess demand indicator and the estimate for the factor of proportionality δ . This is because an estimate for excess demand is given by $\hat{\delta}i_t$, with i_t being the indicator variable which is either equal to $\Delta r_{L,t}$ or z_{t+1} , while an estimate for the change in excess demand is given by $\hat{\delta}\Delta i_t$ – see also the appendix. Due to the importance of the excess demand indicator we have conducted a preliminary analysis in which we compare Δr_L and z . This comparison yields the following results.

Figure 3.2 depicts the levels of the two indicators for the full sample period. The upper half of the graph shows indicator z (right axis) while indicator Δr_L is depicted in the lower half (left axis) – the measurement unit for both axis is percentage points. An indicator level below the zero line implies that credit supply is bigger than credit demand, i.e. excess supply of credit, while an indicator level above the zero line represents excess demand for credit.

The first observation from Figure 3.2 is that both series indicate times of excess credit demand and excess credit supply. However, the indicator Δr_L is much less persistent than z . As a result, the former indicates more regime changes than the latter. Furthermore, the two indicators provide very different information regarding the regime, i.e. demand or supply, that determines the observed credit quantity. For instance, while the change in the loan rate indicates that the credit quantity observed between 2000 and 2002 was mainly determined by demand, implying excess supply of credit, the net percentage of banks tightening credit standards indicates an excess demand for credit – the same is true for the time period of the financial crisis that started in 2007.

Next consider Figure 3.3 and Figure 3.4 which depict first differences of the two indicators. The bars in both graphs represent changes in excess

²²Augmented Dickey-Fuller tests reveal that for none of our time series we can reject the null hypothesis of a unit root at the 1% level of significance except for the indicator variable z . Further tests indicate that the first differences of the series are stationary. See also Figure 3.5 and Figure 3.6 in the appendix for a graphical representation of our data.

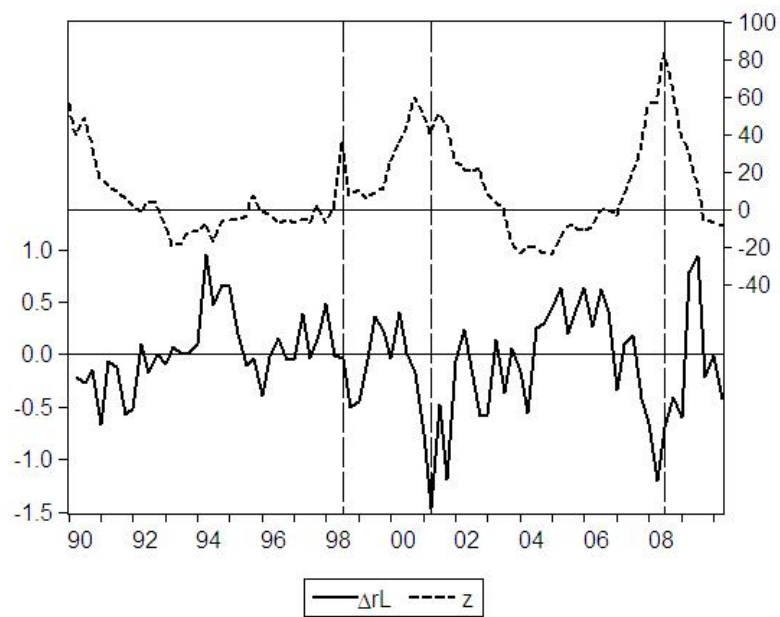


Figure 3.2: Comparison of Indicators (Levels)

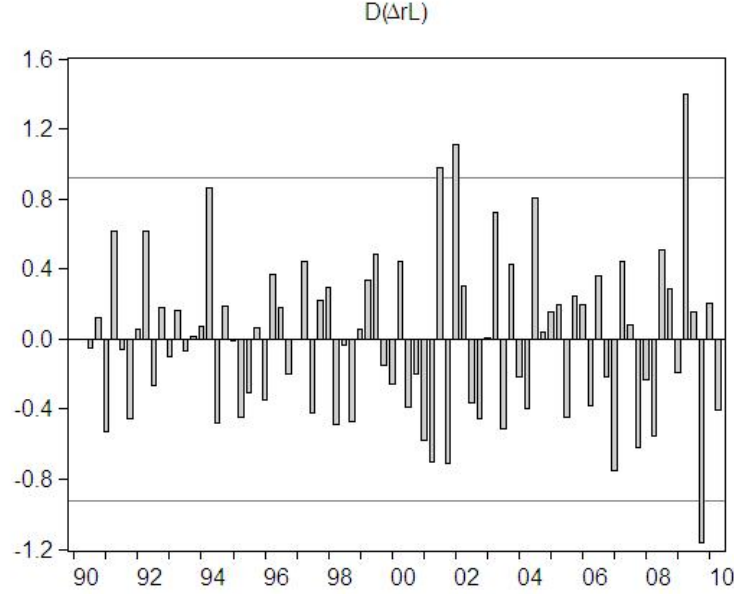


Figure 3.3: Indicator Δr_L (First Differences)

demand while the areas between the horizontal lines indicate the interval given by the mean of the first differences plus/minus two standard deviations.

As one can see from the two figures, most changes in excess demand are within the interval. Thus, both indicators suggest that changes in excess demand are usually not drastic. Furthermore, there is only one time period for which both indicators display a sharp increase in credit rationing, i.e. an increase that exceeds two standard deviations, namely the financial crisis that started in 2007 – note, however, that the quarters in which these sharp increases are indicated are not the same.

Although this is a promising result the discrepancies between the two indicators described above raise the question which of the two provides a more accurate picture of credit rationing in the U.S. market for C&I loans. In section 3.3.1 we have argued that changes in the loan rate are more related to changes in the interest rate target of the FED than to excess credit demand.

More important, though, seems to be the fact that theoretical evidence suggests that in times of high uncertainty or stress credit is likely to be rationed more. The dashed vertical lines in Figure 3.2 are three examples for

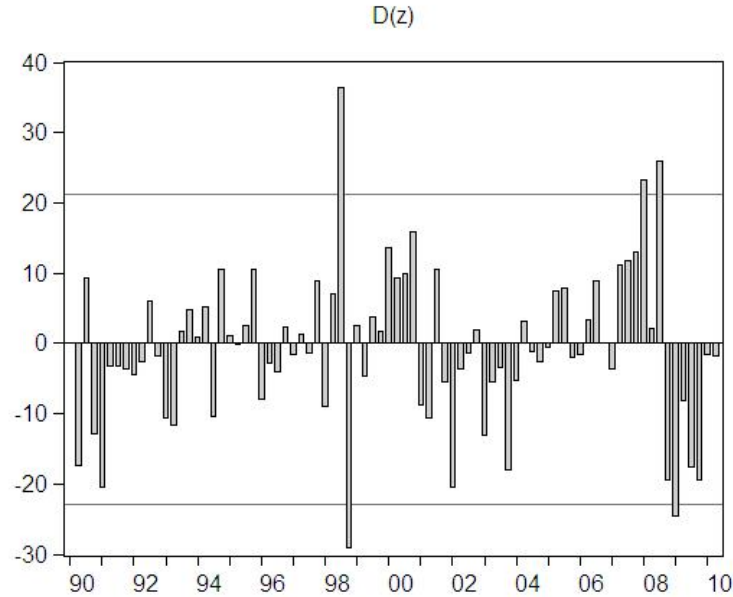


Figure 3.4: Indicator z (First Differences)

such time periods – from left to right they indicate the Russian debt crisis and the collapse of Long Term Capital Management respectively in the third quarter of 1998, the middle of the bear market in the U.S. that followed the bursting of the dotcom bubble and which lasted from the first quarter of 2000 to the fourth quarter of 2002 and finally the bankruptcy of Lehman Brothers in the third quarter of 2008 at the height of the subprime mortgage crisis. As the graph demonstrates, indicator z reaches peaks during all three time periods, i.e. it indicates that credit rationing was relatively high during these events which is consistent with expectations. The change in the loan rate, on the other hand, does not indicate excess demand for credit during any of these events and actually indicates the opposite.²³ The likely reason for this is the aforementioned relation between the loan rate and the federal funds rate and the fact that the FED tried to calm financial markets by lowering

²³The so-called savings and loan crisis, which affected the U.S. economy until the beginning of the 1990ies, is another example in this respect. This time period is generally associated with a credit crunch. While the change in the loan rate does not indicate credit rationing, the net percentage of banks tightening credit does (see Figure 3.2).

the federal funds rate during all these events. Taken together, we interpret these results as evidence in favour of the excess demand indicator z .

The preceding analysis shows that both excess demand indicators suggest that U.S. commercial banks increased non-price rationing of C&I loans sharply during the recent financial crisis. Whether this increase was economically significant or not depends on the estimate for δ which we deal with in the next section.

3.3.5 Estimation Results

In the following, we refer to the disequilibrium model in which the change in the loan rate serves as excess demand indicator as model C. Model E, in turn, in which the loan rate is assumed to be exogenous, is based on the indicator proposed by Harris (1974). Model F employs the same indicator but the loan rate is implicitly endogenous.

Besides estimating our models in first differences, we apply three-stage least squares instead of the conventional system version of two-stage least squares. This is because exogenous shocks are likely to affect both the change in credit demand and the change in credit supply, implying that the disturbance terms in our equations are contemporaneously correlated. Our estimation results are summarised in Table 3.1 and Table 3.2.²⁴ Due to the presence of lagged endogenous variables in all three models, we have to be particularly careful with regard to autocorrelated disturbances. Tests for autocorrelation, however, suggest that we cannot reject the null hypothesis of no autocorrelation up to lag twelve at a standard level of significance in any of our models. Jarque-Bera tests further indicate that the disturbances in all three models are normally distributed.

In terms of the goodness of fit we observe that the variation in the explaining variables accounts for a high fraction of the total variation in the explained variable. In addition, we find that the fit for the equation describing the change in supply is higher in all three models. This could reflect the fact that the demand specification is less grounded in theory than the supply specification.

Examining the parameter estimates we see that the coefficient on the change in capacity utilisation ΔCU_{t-1} has the wrong sign and is insignificant

²⁴Numbers in parenthesis are t-statistics while the stars indicate the level of significance, i.e. * implies significance at the 10% level, ** indicates significance at the 5% level and *** stands for the 1% level of significance.

Equation:			
$\Delta q_t = \beta_1 + \beta_2 \Delta(r_{L,t} - r_{T,t}) + \beta_3 \Delta TD_t + \beta_4 \Delta IP_{t-1} + \beta_5 \Delta CD_t + \beta_6 (\Delta q_{t-1} + \delta \Delta i_{t-1}^-) - \delta \Delta i_t^- + u_{s,t}$			
Coeff. Estim.	Model C	Model E	Model F
<i>const.</i>	-2.5848 (-1.9277)*	-2.6300 (-2.2594)**	-2.5404 (-1.6339)
$\Delta(r_{L,t} - r_{T,t})$	26.8972 (2.6670)***	8.6613 (2.6785)***	32.5792 (2.8439)***
ΔTD_t	0.0378 (2.2226)**	0.0352 (2.7297)***	0.0407 (1.9386)*
ΔIP_{t-1}	4.6291 (3.9912)***	5.2967 (5.5395)***	4.0523 (2.9552)***
ΔCD_t	5.0063 (1.9493)*	5.4404 (2.9111)***	6.9982 (2.4528)**
Δq_{t-1}	0.7705 (16.7215)***	0.7665 (17.2122)***	0.7648 (14.3948)***
Δi_t^-	2.7121 (1.4553)	0.1129 (1.4424)	0.2377 (2.2611)**
Obs.	79	80	80
adj. R^2	0.8151	0.8162	0.7976

Table 3.1: Estimation Results (Change in Credit Supply)

Equation:			
$\Delta q_t = \alpha_1 + \alpha_2 \Delta(r_{A,t} - r_{L,t}) + \alpha_3 \Delta PT_{t-1} + \alpha_4 \Delta IP_{t-1} + \alpha_5 \Delta CU_{t-1} + \alpha_6 (\Delta q_{t-1} + \delta \Delta i_{t-1}^+) - \delta \Delta i_t^+ + u_{d,t}$			
Coeff. Estim.	Model C	Model E	Model F
<i>const.</i>	-2.3578 (-1.4480)	-1.7630 (-1.2142)	-2.5963 (-1.5622)
$\Delta(r_{A,t} - r_{L,t})$	-0.5548 (-0.2198)	-3.4126 (-1.7285)*	-0.4345 (-0.1596)
ΔPT_{t-1}	0.0762 (2.8841)***	0.0420 (2.1035)**	0.0914 (3.3632)***
ΔIP_{t-1}	5.4412 (2.3727)**	5.0613 (2.6604)***	5.7729 (2.3798)**
ΔCU_{t-1}	-1.6360 (-0.6100)	-0.7685 (-0.3730)	-1.6816 (-0.5932)
Δq_{t-1}	0.8401 (13.8968)***	0.8164 (14.6840)***	0.8510 (14.7503)***
Δi_t^+	2.7121 (1.4553)	0.1129 (1.4424)	0.2377 (2.2611)**
Obs.	79	80	80
adj. R²	0.7585	0.7564	0.7690

Table 3.2: Estimation Results (Change in Credit Demand)

in all three models. The coefficient on the change in the lagged index of industrial production ΔIP_{t-1} , on the other hand, is significantly different from zero in both equations and in all three models. Besides, it has the right sign, i.e. an increase in the index, indicating a more favourable economic outlook, increases both credit demand and credit supply. Since capacity utilisation is more related to long term investments these results could indicate that firms use bank credit mainly to finance their working capital.

Next consider the coefficient on the change in the spread between the rate on alternative forms of external financing and the loan rate, i.e. $\Delta(r_{A,t} - r_{L,t})$. Although the coefficient has the right sign in all three models, meaning that an increase in the spread causes credit demand to increase, it is significantly different from zero only in model E. We believe that the insignificance is either due to the fact that the interest rate on AAA-rated corporate bonds is a bad proxy for the rate on alternative external funds or that credit demand does not depend on the spread but the loan rate itself. The main intuition for the latter is that, as we have pointed out above, the only source of external financing for small and medium sized firms is bank credit. Therefore credit demand does not depend on the loan rate relative to the rate on other sources of funds but on the level of the loan rate.²⁵

The coefficient on the change in the spread between the loan rate and the rate on alternative assets, i.e. $\Delta(r_{L,t} - r_{T,t})$, in turn, is significantly different from zero and has the right sign in all three models. Hence, if the spread increases credit supply increases too.

Furthermore we find that the coefficients on ΔTD_t and ΔPT_{t-1} , i.e. on the change in the amount of total deposits and the change in profits after taxes lagged by one period, are significant in all three models. The coefficient sign of ΔTD_t is positive, indicating that an increase in the amount of loanable funds implies that credit supply increases. The positive sign of the coefficient on the change in profits supports the hypothesis that it is easier for profitable firms to get external financing and that these firms do not substitute external with internal funds.

²⁵Using the change in the loan rate instead of the change in the spread does not change the results dramatically. In fact they get worse – the coefficient on ΔTD_t , for instance, becomes insignificant in model C. Unfortunately we could not check whether we have a bad proxy due to the absence of any alternative time series for the rate on alternative forms of external financing: while the series on the commercial paper rate is too short, the series on the rate on BAA-rated corporate bonds is nearly identical with the series for AAA-rated bonds.

The coefficient on the change in the costs per dollar of deposits ΔCD_t is found to be significantly different from zero in all three models. The positive sign of the coefficient indicates that credit supply increases if the costs per dollar of deposits increase. This counterintuitive result lends support to the hypothesis that if banks rely more on time and savings deposits they need less liquid assets and can therefore invest their funds in more illiquid assets such as loans.

Finally, the change in the lagged quantity of loans outstanding Δq_{t-1} is highly significant in both equations and in all three models. The magnitude of these estimates suggests that there are high adjustment costs to the quantity of outstanding credit for both lenders and borrowers, although they seem to be higher for borrowers.

As already indicated above, however, the coefficient estimate we are most interested in is $\hat{\delta}$, i.e. the coefficient on the (change in the) indicator variable. This estimate allows us to calculate an estimate for (the change in) excess credit demand. In model C the coefficient has the right sign but is not significantly different from zero. According to Quandt (1988) this result indicates that the speed of adjustment of the loan rate in the U.S. market for C&I loans to imbalances between credit supply and demand is very high, meaning that the equilibrium framework would be more appropriate. This finding is consistent with the conclusions of Ito & Ueda (1981) and Kugler (1987). In model E, in which we use the indicator proposed by Harris (1974), the estimate for δ is also not significantly different from zero. However, if we relax the assumption of an exogenous loan rate, as in model F, the coefficient becomes significant.²⁶ Nevertheless, the estimated coefficient appears to be rather small, i.e. $\hat{\delta} = 0.2377$.

Based on this parameter estimate we have calculated estimates for credit demand and supply respectively, excess demand, the change in excess demand and the ratio of excess demand to total demand. Table 3.3 shows our results for the time period from the first quarter of 2007 to the fourth quarter of 2009. From the table we conclude the following.

The second largest increase in excess credit demand in the whole sample period, i.e. from the first quarter of 1990 to the second quarter of 2010, namely 6.18 bn (constant 2005 USD) occurred from the second to the third

²⁶If we include the residuals from the reduced form estimation of the change in the loan rate as explanatory variable in both structural equations we find that the estimated coefficients on this residual variable are significantly different from zero – at least in the supply equation. Therefore we conclude that the loan rate is in fact endogenous.

Quarter	\hat{d}	\hat{s}	$\hat{d} - \hat{s}$	$\Delta(\hat{d} - \hat{s})$	$(\hat{d} - \hat{s})/\hat{d}$
2007:q1	949.74	950.62	-0.88	-0.88	–
2007:q2	975.51	973.72	1.78	2.66	0.18%
2007:q3	1,024.05	1,019.48	4.56	2.78	0.45%
2007:q4	1,081.93	1,074.27	7.65	3.09	0.71%
2008:q1	1,127.19	1,114.02	13.17	5.52	1.17%
2008:q2	1,133.41	1,119.72	13.69	0.52	1.21%
2008:q3	1,141.79	1,121.91	19.87	6.18	1.74%
2008:q4	1,170.31	1,155.05	15.26	-4.61	1.30%
2009:q1	1,124.20	1,114.79	9.41	-5.85	0.84%
2009:q2	1,083.99	1,076.51	7.45	-1.93	0.69%
2009:q3	1,012.06	1,008.73	3.33	-4.16	0.33%
2009:q4	955.29	956.60	-1.31	-4.64	–

Table 3.3: Credit Market Estimates

quarter of 2008. This finding suggests that we cannot reject our working hypothesis, namely that U.S. commercial banks increased non-price rationing sharply during the crisis. Furthermore, at the height of the financial turmoil, i.e. in the third quarter of 2008 when the failure of Lehman Brothers shook the global financial system to its very foundations, excess demand for credit reached its highest level in the whole sample period, totalling 19.87 bn (constant 2005 USD) – this figure corresponds to 1.74% of total demand for credit. In other words, the level of bank debt firms wanted to hold was 19.87 bn higher than what commercial banks were willing to offer.

However, one has to put these numbers into perspective by considering information from the Survey of Terms of Business Lending.²⁷ According to survey data the total value of C&I loans issued by domestically chartered commercial banks in the U.S. in the first full business week in November 2008, i.e. shortly after the collapse of Lehman Brothers, was 45 bn (constant 2005 USD). Thus, the estimated extent of excess demand for credit is negligible small.

Together, these results suggest that although there is credit rationing in

²⁷In the Survey of Terms of Business Lending the U.S. Federal Reserve collects information concerning price and non-price terms of business loans made during the first full business week of the mid-month of each quarter. The collected data also include detailed information on individual business loans made during the survey week. Based on this information monetary authorities estimate the terms of business lending in the whole U.S. market for C&I loans. For more information see <http://www.federalreserve.gov/releases/e2/>.

the U.S. market for C&I loans and there was an increase in rationing during the recent financial crisis, excess demand for credit does not constitute an important phenomenon. In particular, the low degree of rationing casts into doubt the conjecture that excess credit demand had a substantial effect on the real sector. This conclusion is also consistent with the latest findings of Berger & Udell (1992) presented in section 3.2.1.

3.4 Conclusion

In this chapter we address the question whether U.S. commercial banks increased non-price credit rationing sharply in the market for C&I loans during the financial crisis that started in 2007. For this purpose we unify the disequilibrium econometric framework with the survey approach to credit rationing. This combined approach not only employs an excess demand indicator that is likely to be superior to previous indicators but it also allows us to account for many potential sources of misspecification due to the tractability of disequilibrium models with sample separation information. While a preliminary analysis shows that there was indeed a sharp increase in credit rationing during the crisis, our estimation results suggest that credit rationing in general is only a minor issue. In particular, our findings do not support the popular view that non-price credit rationing acted as a major transmission channel during the crisis. However, our analysis does not rule out the possibility that conventional price rationing played an important role in terms of the transmission of the crisis to the real sector. Furthermore, we do not exclude that banks resorted to non-price credit rationing in other loan markets such as the wholesale funds markets, in which banks lend excess funds to other banks.

3.5 Appendix

Consider the following (simplified) dynamic disequilibrium model:

$$d_t = \alpha_1 + \alpha_2 p_t + \alpha_3 x_{d,t} + \alpha_4 d_{t-1} + \epsilon_{d,t} \quad (3.8)$$

$$s_t = \beta_1 + \beta_2 p_t + \beta_3 x_{s,t} + \beta_4 s_{t-1} + \epsilon_{s,t} \quad (3.9)$$

$$q_t = \min(d_t, s_t) \quad (3.10)$$

$$d_t - s_t = \delta i_t \quad (3.11)$$

Note that for $i_t = \Delta p_t$ and $\delta = 1/\gamma$ a dynamic version of model C is obtained while for $i_t = z_{t+1}$ sample separation information is based on the excess demand indicator proposed by Harris (1974). From the minimum condition (3.10) and the deterministic indicator function (3.11) one can infer that in case of $d_t < s_t$ demand and supply are equal to $d_t = q_t$ and $s_t = q_t - \delta i_t$ respectively while if $d_t \geq s_t$ demand and supply are given by $d_t = q_t + \delta i_t$ and $s_t = q_t$.

In the second case, the system can be written as

$$\begin{aligned} q_t + \delta i_t &= \alpha_1 + \alpha_2 p_t + \alpha_3 x_{d,t} + \alpha_4 d_{t-1} + \epsilon_{d,t} \\ q_t &= \beta_1 + \beta_2 p_t + \beta_3 x_{s,t} + \beta_4 s_{t-1} + \epsilon_{s,t} \end{aligned}$$

while in the first it is given by

$$\begin{aligned} q_t &= \alpha_1 + \alpha_2 p_t + \alpha_3 x_{d,t} + \alpha_4 d_{t-1} + \epsilon_{d,t} \\ q_t - \delta i_t &= \beta_1 + \beta_2 p_t + \beta_3 x_{s,t} + \beta_4 s_{t-1} + \epsilon_{s,t} \end{aligned}$$

Using the two auxiliary variables

$$\begin{aligned} i_t^+ &= \begin{cases} i_t & \text{if } i_t > 0 \\ 0 & \text{otherwise} \end{cases} \\ i_t^- &= \begin{cases} -i_t & \text{if } i_t \leq 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

the two cases from above can be combined to yield

$$\begin{aligned}
q_t + \delta i_t^+ &= \alpha_1 + \alpha_2 p_t + \alpha_3 x_{d,t} + \alpha_4 d_{t-1} + \epsilon_{d,t} \\
q_t + \delta i_t^- &= \beta_1 + \beta_2 p_t + \beta_3 x_{s,t} + \beta_4 s_{t-1} + \epsilon_{s,t}
\end{aligned}$$

The (unobserved) lagged endogenous explanatory variables d_{t-1} and s_{t-1} are replaced in a similar way, noting that equation (3.11) holds for all periods. This finally yields the system

$$q_t + \delta i_t^+ = \alpha_1 + \alpha_2 p_t + \alpha_3 x_{d,t} + \alpha_4 (q_{t-1} + \delta i_{t-1}^+) + \epsilon_{d,t} \quad (3.12)$$

$$q_t + \delta i_t^- = \beta_1 + \beta_2 p_t + \beta_3 x_{s,t} + \beta_4 (q_{t-1} + \delta i_{t-1}^-) + \epsilon_{s,t} \quad (3.13)$$

If the indicator variable is the change in the price, such as in model C, then the endogenous variables of the transformed system are q_t , p_t , i_t^+ and i_t^- . On the other hand, if indicator z is used then the endogenous variables are q_t , i_t^+ and i_t^- . In the latter case, p_t would be exogenous as in model A. Assuming, however, that the price adjusts to imbalances between demand and supply, we can treat p_t as implicitly endogenous in the estimation of equations (3.12) and (3.13) as suggested by Rudebusch (1987). The major advantage of this procedure is that it does not require an explicit specification of the price adjustment equation.

Further note that from equation (3.11) it is apparent that an estimate for excess demand is given by $\hat{\delta} i_t$, where $\hat{\delta}$ is an estimate for the true parameter value. Besides, first differencing the indicator equation (3.11) yields $\Delta d_t - \Delta s_t = \delta \Delta i_t$. An estimate for the change in excess demand is then given by $\hat{\delta} \Delta i_t$ because $\Delta d_t - \Delta s_t = (d_t - s_t) - (d_{t-1} - s_{t-1})$.

3.5.1 Data Description

Variable	Description	Source
q	Commercial and Industrial Loans Outstanding (domestically chartered commercial banks in the U.S.), 3-month averages, seasonally adjusted, billions of constant 2005 dollars	Assets and Liabilities of Commercial Banks in the U.S. (H.8), Federal Reserve
r_L	Bank Prime Loan Rate (p.a.), 3-month averages, adjusted for annual inflation	Selected Interest Rates (H.15), Federal Reserve
r_T	Interest Rate on 6-month U.S. Government Treasury Bills (p.a.), 3-month averages, adjusted for annual inflation	Selected Interest Rates (H.15), Federal Reserve
TD	Total Deposits (domestically chartered commercial banks in the U.S.), 3-month averages, seasonally adjusted, billions of constant 2005 dollars	Assets and Liabilities of Commercial Banks in the U.S. (H.8), Federal Reserve
CD	Costs Per Dollar of Deposits, ratio of time and savings deposits to total deposits time the 6-month interest rate on certificates of deposits (p.a.)	Flow of Funds Account of the U.S. (Z.1), Federal Reserve
IP	Index of Industrial Production, total index (including all major industry groups), seasonally adjusted	Industrial Production and Capacity Utilization (G.17), Federal Reserve
r_A	Interest Rate on AAA Corporate Bonds (p.a.), 3-month averages, adjusted for annual inflation	Selected Interest Rates (H.15), Federal Reserve
PT	Profit after Tax (nonfarm nonfinancial corporate business in the U.S.), seasonally adjusted, billions of constant 2005 dollars	Flow of Funds Account of the U.S. (Z.1), Federal Reserve
CU	Capacity Utilisation, total index, seasonally adjusted	Industrial Production and Capacity Utilization (G.17), Federal Reserve
z	Net Percentage of Banks Reporting a Tightening of Credit Standards (for large and medium sized firms)	Senior Loan Officer Opinion Survey on Bank Lending Practices, Federal Reserve

Table 3.4: Description of Data

3.5.2 Graphical Data Representation

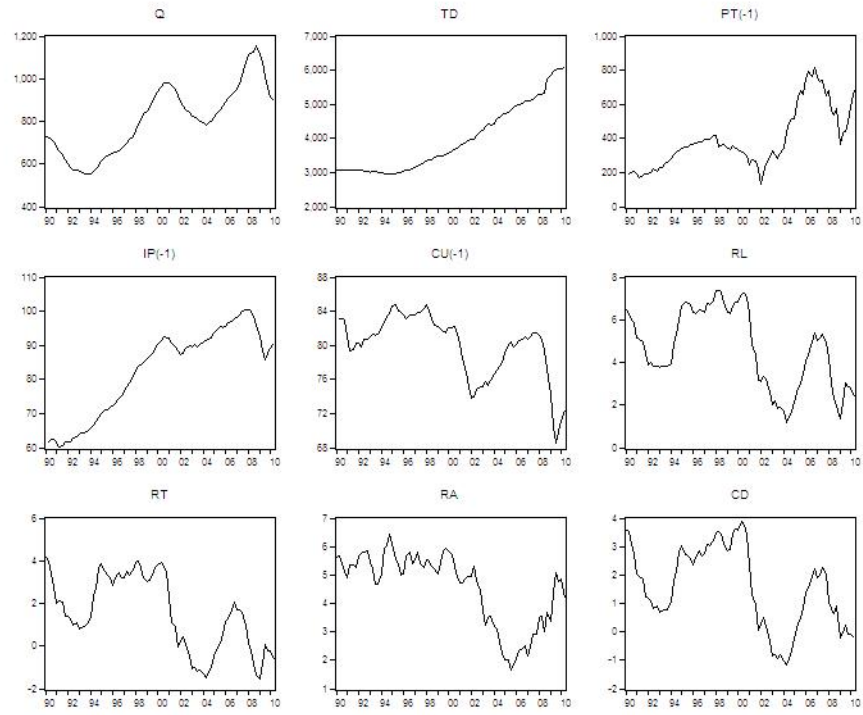


Figure 3.5: Level Data

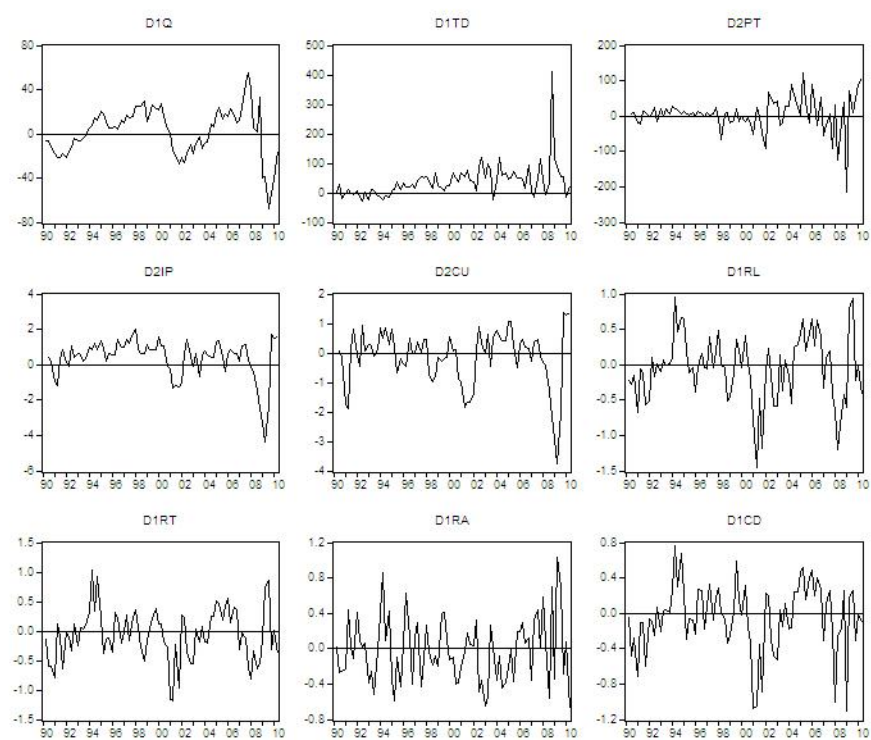


Figure 3.6: Data in First Differences

Chapter 4

Wholesale Funding and Commercial Bank Lending – Implications from the 2008 Liquidity Crunch

4.1 Introduction

Maturity transformation, i.e. converting short-term liabilities such as retail deposits into long-term assets, is one of the major functions performed by financial intermediaries. The resulting maturity mismatch, however, leaves intermediaries like banks particularly vulnerable to runs by uninformed depositors.¹ For this reason many economies have introduced deposit insurance so that traditional bank runs have become an outdated phenomenon.

In the recent past, though, banks have increasingly supplemented their deposit base with so-called wholesale funds which are often provided by arm's length markets. During the global financial crisis that started in 2007 these markets essentially shut down. The sudden dry up of wholesale funding sources combined with the increased reliance on these funds resulted in the modern equivalent of a bank run. Unfortunately, the modern form, which has been referred to as liquidity crunch by Brunnermeier (2009), unfolded similar deleterious effects as its traditional counterpart. In fact, it not only amplified the crisis – which was thought to be confined to the subprime

¹More information on traditional bank runs is provided in Diamond & Dybvig (1983).

mortgage sector – but it also helped to propagate it.

While scholars have mainly emphasised the advantages of bank wholesale funding, the crisis has provoked a shift towards the disadvantages. The empirical analysis presented in this chapter follows this development. In particular, we study whether the liquidity crunch negatively affected the lending behaviour of U.S. commercial banks that relied more heavily on wholesale funds. Although our analysis is closely related to the study by Ivashina & Scharfstein (2010), there are important differences. First of all, we use a much larger dataset on individual bank level data that covers all insured U.S. commercial banks. Besides, the dataset not only captures large syndicated loans but various types of loans. Furthermore, we control for other factors that may have caused a change in the supply of credit, such as equity capital holdings. Finally, and probably most important, we analyse whether the crunch had a larger effect in terms of lending on banks with less liquid balance sheets. To our knowledge, the relation between wholesale dependence and liquidity during a crisis has not been analysed before.

Our results are particularly interesting with regard to banking regulation. For instance, the Basel Committee on Banking Supervision has suggested the introduction of minimum global liquidity standards in response to the observed run by lenders.² The objective of these standards is twofold: first to strengthen the resilience of banks to short-term disruptions in the access to funding and second to provide incentives for banks to use more stable sources of funding. Our analysis, in turn, allows us to evaluate whether the introduction of a short-term liquidity coverage ratio is likely to mitigate the effects of a disruption in funding access. Furthermore, if lending is found to be strongly affected by the reliance on wholesale funds – potentially having negative effects on the real economy – then the committee’s proposal to curtail the use of wholesale funds and foster the use of more stable sources of funds receives further support.

The first part of this chapter, consisting of sections 4.2.1 to 4.2.3, discusses bank wholesale funding and provides some details on the liquidity crunch and its implications for bank lending. Section 4.3, which is the main part of this chapter, presents our empirical analysis, i.e. the specifications we estimate, the dataset we use and the results of our estimations. Finally, section 4.4 concludes.

²See, for instance, *The Basel Committee’s response to the financial crisis: report to the G20*, published by the Bank of International Settlement in October 2010.

4.2 Review

4.2.1 Bank Wholesale Funding

Banks cannot raise an infinite amount of funds via retail deposits. This is because they can only access a limited pool of potential depositors.³ Therefore banks also use funds from various other sources. These additional funds are typically provided by professional investors, banks and other corporations which is why they are called wholesale funds, managed liabilities or non-retail funds. Banks usually raise these funds on a short-term rollover basis, i.e. they continuously tap the wholesale funds markets to repay maturing liabilities. One reason for this is that funds borrowed long-term are associated with higher costs. The instruments used for this type of short-term funding include certificates of deposits or brokered deposits, repurchase agreements, Fed funds, interbank loans and commercial paper. In contrast to retail deposits, wholesale funds are in most instances uninsured and they are not subject to a reserve requirement. Furthermore, they are mostly provided by arm's length markets implying that their availability crucially depends on the smooth functioning of these markets.

Albeit wholesale funding is not a new phenomenon in banking, it has attracted more and more interest in the recent past. This is mainly due to the fact that banks increasingly use wholesale funds to supplement their deposit base. Table 4.1 illustrates this point: while the fraction of deposits to assets declined by more than ten percentage points in the time period from 1980 to 2010, the fraction of all other, non-retail liabilities increased.⁴ Thus, Table 4.1 supports the conjecture that there has been a noticeable change in the way how banks finance their assets. The advantages and disadvantages respectively this change entails are the subject of a relatively young research literature whose findings we briefly summarise below.

From a social point of view, probably the most important argument for wholesale funding is that banks can finance socially beneficial investments without being constrained by a local deposit base. According to Calomiris (1999), another benefit is that wholesale financiers provide for market dis-

³Despite globalisation and the Internet revolution, retail banking can still be characterised as a local business. This is probably due to high transaction costs.

⁴Table 4.1 is based on balance sheet data from the Consolidated Reports of Condition and Income, the so-called Call Reports. They are collected quarterly and cover all insured U.S. commercial banks. For more information on our dataset see section 4.3.2.

	1980q1	1985q1	1990q1	1995q1	2000q1	2005q1	2010q1
# of Banks	14,362	14,364	12,441	10,151	8,439	7,517	6,703
<i>% of Total Assets</i>							
Total Deposits	80.27	77.96	76.60	69.53	66.57	66.60	68.78
Federal Funds Borrowed	6.52	7.61	8.49	7.80	8.18	6.92	4.72
Subordinated Debt	0.36	0.47	0.59	1.00	1.37	1.35	1.25
Equity	5.88	6.29	6.37	7.86	8.40	10.05	10.72
Other Liabilities	6.97	7.68	7.94	13.81	15.48	15.08	14.53

Table 4.1: Liability Structure of U.S. Commercial Banks

cipline of banks by threatening to withdraw (uninsured) funds if the assets of a bank become too risky. In a similar vein, Feldman & Schmidt (2001) argue that prices paid by banks on uninsured wholesale deposits could help regulators to assess the financial condition of banks. In particular, higher market prices would send a signal of greater bank risk taking to regulators. From the perspective of banks, on the other hand, wholesale funds represent a relatively attractive way to expand their operations, i.e. to increase the size of their balance sheets. Moreover, the higher leverage implies that banks can increase their profitability measured as the return on equity (ROE), for instance. Mishkin (2004) further notes that these funds can be used to refinance unexpected retail withdrawals.

However, there is also a dark side of bank wholesale funding, as emphasised by Huang & Ratnovski (2011). Their paper challenges the assumption of Calomiris (1999) that wholesale financiers have an incentive to monitor the risk taking of individuals banks. In particular, they show that if there is a costless but noisy public signal on the quality of a bank's assets, financiers have a lower incentive to monitor the bank. Under certain conditions this can lead to a situation in which wholesale financiers withdraw their funds based on an erroneous signal, implying that a healthy bank faces a run by lenders – as we argue below, this is likely what happened during the recent financial crisis. Another disadvantage of wholesale funding was identified by Kashyap & Stein (2000) in a paper on the bank lending channel. They show that small banks, which presumably have less access to wholesale funds markets, react stronger to a monetary contraction in terms of lending than big banks with better access to wholesale funds. Thus, the effectiveness of monetary policy could be reduced by an increased use of wholesale funding.

The financial crisis that started in 2007 has revealed yet another weakness of short-term funding via arm's length markets. One result of the turmoil in these markets, equatable only with a system-wide bank run, was that a

vicious deleveraging process was set in motion which not only amplified the crisis but also helped to propagate it. Since the events during and after this modern equivalent of a bank run, which has been referred to as liquidity crunch, are most important for our analysis, we discuss them in more detail in the following section.

4.2.2 The Liquidity Crunch

The term liquidity is used in different ways. One definition that suits our purposes comes from Brunnermeier & Pedersen (2009) who distinguish between market liquidity and funding liquidity. While the former is characterised as the ease with which an asset can be traded, the latter is the ease with which investors can obtain funding. In other words, at higher funding liquidity it is easier for investors to raise money in the market. Drehmann & Nikolaou (2009), in turn, offer a related definition in which funding liquidity is the ability to settle obligations immediately when they become due. Based on this definition it is straightforward to define the concept of funding liquidity risk as the risk that over a specific time horizon a bank will be unable to settle its obligations immediately.⁵

According to Brunnermeier (2009) funding liquidity risk is made up of three types of risk. The first is margin or haircut risk. In a repurchase agreement, for instance, a bank sells securities and agrees to buy back these securities at some point in the future. In such a transaction the bank usually does not get the total value of the securities it sells and the (percentage) difference between the value of the securities and the amount it gets is called haircut. If the haircut increases, a bank can raise fewer funds from selling securities temporarily. Margin risk, on the other hand, is associated with margin lending where a bank borrows funds to acquire an asset and where it uses the asset as collateral. Typically, banks cannot borrow the whole amount and the difference is called margin. Since the margin is not fixed, banks face the risk of a margin call, i.e. they may have to provide additional funds. The second risk is rollover risk. As already outlined above, banks continuously tap wholesale funds markets, like the commercial paper market or the interbank market, to repay maturing liabilities, i.e. they roll over their debt. If investors buy less commercial paper or banks are reluctant

⁵Notice that in the definition of Drehmann & Nikolaou (2009) funding liquidity is a point-in-time, zero-one concept while funding liquidity risk is forward looking and depends on a specific time horizon. Besides, the latter can take on many values.

to make loans to other banks then rolling over debt becomes more difficult. And finally there is the risk of withdrawal, meaning that wholesale financiers, such as hedge funds, who provide funds via large deposits, suddenly withdraw their funds.

During the recent financial crisis funding liquidity evaporated in wholesale funds markets because all three risks materialised. Shin (2009), for instance, concludes that the failure of the British savings and mortgage bank Northern Rock can be mainly attributed to the fact that this highly leveraged bank was not able to roll over its maturing debt in the commercial paper market. In contrast, according to Brunnermeier (2009) an important factor in the demise of the investment bank Bear Stearns was that in March 2008 the bank was suddenly unable to secure funding in the repo market due to rumours about the solvency of the bank. The situation was further aggravated because hedge fund clients of the bank withdrew funds on a large scale.

Figure 4.1, which depicts a popular indicator for the condition of wholesale funds markets, provides further evidence for the dry up in wholesale funding sources and the subsequent liquidity crunch. It shows the spread between the three-month London Interbank Offered Rate (3m-Libor) for U.S. Dollar denominated loans and the rate on three-month U.S. Treasury Bills between 2006 and 2010. The graph indicates that the spread, also referred to as TED spread, tends to be quite low implying that the risk of lending to a bank instead of lending to the U.S. government is perceived to be small. In other words, banks can obtain funds relatively cheap in the interbank market. With the onset of the financial crisis in 2007, however, the costs of borrowing in the interbank market rose sharply, reaching an all-time high in September 2008 when the investment bank Lehman Brothers declared bankruptcy.

The examples mentioned above highlight the fact that it became increasingly difficult for banks to settle their obligations due by using wholesale funds. This has led Brunnermeier (2009) to refer to the worsening liquidity situation as liquidity crunch. The question is, how this crunch fits into the wider context of the financial crisis. Blanchard (2008), for instance, characterises its role during the recent crisis as follows.

The bursting of the bubble in the U.S. housing market in early 2007 resulted in great uncertainty about the value of a large number of assets, including highly complex and opaque mortgage related credit derivatives, which had been distributed around the globe as a result of securitisation. This not only increased the uncertainty about the value of bank capital but also the probability of insolvency of many financial institutions. Wholesale

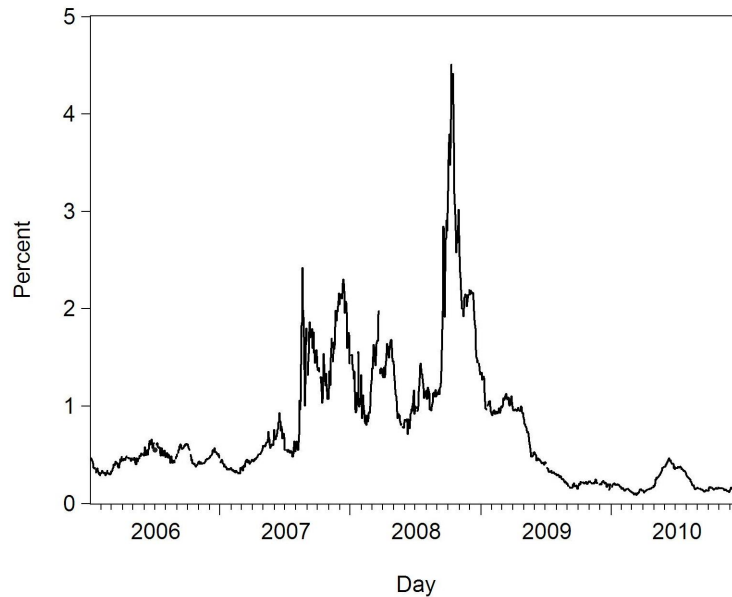


Figure 4.1: TED Spread

financiers, who had become more and more important, therefore wanted to take their funds out of institutions that they perceived to be too risky. Depositors, on the other hand, had no incentive to take out their funds since they were mostly insured. The run by lenders was not targeted at specific banks as in a traditional bank run but this time it was system-wide. As a result, banks could no longer refinance themselves – or only at very high costs – via wholesale funds markets. In addition, Ivashina & Scharfstein (2010) note that the run by lenders was accompanied by a run of borrowers who drew down on existing credit lines for fear of a shortage of bank credit, thereby aggravating the funding situation of banks even further.

This liquidity crunch had the same consequences as a traditional run: since banks could not borrow funds they had to shrink the asset side of their balance sheets, i.e. they had to deleverage – for a detailed analysis of the deleveraging process see section 4.2.3. Furthermore, banks had to maintain adequate capital ratios for regulatory reasons, for instance. Since additional equity was difficult to obtain in these market circumstances banks had to deleverage even more. This process involved selling assets at fire-sale prices, mainly because there were few deep pocket investors who could counter the excess supply of assets. Therefore asset prices got even more depressed and

the uncertainty about bank capital further increased.

This stylised description of the various stages during the crisis makes clear that a vicious circle was set in motion in which the liquidity crunch played a prominent role. Factors that exacerbated the process include the high leverage of a large number of banks, the complexity of many assets and the shift towards the originate-to-distribute business model which was facilitated by securitisation. Brunnermeier (2009) also emphasises that it was the interrelation between funding and market liquidity that fuelled this vicious circle which he calls the liquidity spiral. However, we are not so much interested in the liquidity crunch and its causes but rather in its implications, which we deal with in the following section.

4.2.3 Wholesale Funding, Liquidity and Bank Lending

In order to analyse the implications of the liquidity crunch we start with a typical balance sheet of a bank.⁶ The right side of this balance sheet shows all bank liabilities, including equity capital, or the sources of funds. The left side, in turn, lists all assets of a bank or the use of funds. Given this snap-shot of the condition of a bank the question is how a liquidity crunch affects bank (lending) behaviour.

Suppose, for instance, that the bank under consideration uses an amount x of wholesale funds to finance a fraction of its assets. For simplicity, further assume that these debts all mature at the same time and are short-term. After a shock to the financial system the bank is not able to roll over the total amount of x in wholesale funds markets but only an amount $y < x$. As a result, the bank is short of funds since in order to settle its obligations equal to x it needs an amount of $(x - y)$ in additional funds.⁷

In such a strained liquidity situation, the bank basically has two options to raise the additional funds needed. First, it could borrow from the lender of last resort, i.e. the central bank. However, as was the case during the recent financial crisis, banks may be reluctant to use this option because they could get stigmatised by the market and thus would even have a harder time to raise funds. Therefore, many banks that have a lack of funding liquidity during a modern bank run probably revert to the second measure, namely

⁶See Mishkin (2004), chapter 9, for an introduction to bank balance sheets.

⁷Notice that the case we describe is similar to a situation in which a bank faces a massive deposit outflow which it cannot cover with its reserves, as in a traditional bank run.

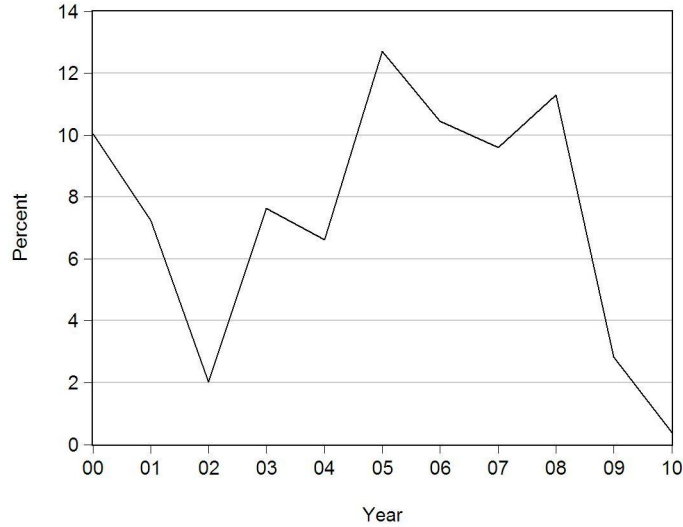


Figure 4.2: Growth Rate of Total Loans Outstanding

to shrink the asset side of the balance sheet to use the proceeds to settle their obligations. This process is referred to as deleveraging and it implies the shrinking of the bank balance sheet to pay off debt. Essentially, a bank can deleverage by selling some of its assets, like bonds or stocks, or it can stop lending – in reality, banks probably implement a combination of both measures. The latter involves that the bank does not make any new loans, it does not roll over existing debt or even calls in loans. Figure 4.2, which is based on Call Report data, suggests that this is indeed what happened during the recent crisis. Although the growth rate of the amount of total loans outstanding increased during the first phase of the crisis from 2007 to 2008, which is probably due to the aforementioned run by borrowers, in the second phase the growth rate declined sharply, implying that banks reduced lending by a considerable amount.

Given these basic balance sheet mechanics, consider a case with two almost identical banks. The only difference between the two is that bank A uses an amount x_A of wholesale funds while bank B uses an amount $x_B < x_A$. Again assume that after a financial shock in form of a liquidity crunch both banks are only able to roll over part of their wholesale debts, i.e. an amount $y < x_B < x_A$. Since bank A needs more additional funds than bank B to settle its obligations, it is reasonable to assume that bank A will deleverage more than bank B , implying that it will also lend less. Therefore our

first hypothesis with regard to the recent financial crisis reads as follows: *U.S. commercial banks which had been more dependent on wholesale funds decreased loan supply more than banks which had been less dependent.*

The motivation for our second hypothesis is based on a different observation. Mishkin (2004), for instance, notes with regard to the deleveraging process that the costliest way for a bank to acquire additional funds is to reduce lending, mainly because the bank loses customers. He argues that selling assets with a high market liquidity is associated with lower costs.⁸ This suggests that a bank with a very liquid balance sheet, meaning that the market liquidity of its assets is very high, probably deleverages by selling assets instead of reducing its lending.⁹ Thus, our second hypothesis can be stated in the following way: *Wholesale dependent U.S. commercial banks which had had a highly liquid balance sheet decreased loan supply less than wholesale dependent banks with less liquid balance sheets.*

The empirical evidence gathered so far suggests that we cannot reject either hypothesis, although for hypothesis number two it is very sparse. Radatz (2010), for instance, uses data for more than 600 banks in over 40 countries, excluding the U.S., to study the propagation of the financial crisis that started in 2007 and its real consequences. With regard to the latter, he analyses the effect of wholesale dependence on bank lending in a simple regression framework. His results indicate that banks with a higher wholesale dependence reduced lending more during the crisis than banks with a lower dependence. He also claims that this relationship is meaningful from an economic perspective.

Another interesting analysis with regard to hypothesis number one is the study by Ivashina & Scharfstein (2010). They too analyse the effect of the modern form of a bank run on lending. Their findings suggest that banks which were more vulnerable to the run, i.e. banks with a higher dependence on wholesale funds and banks which potentially faced more credit line drawdowns, did indeed reduce loan supply more than less vulnerable banks. Unfortunately, their analysis is based on syndicated loans only, implying that it does not cover the large fraction of small non-syndicated loans – another shortcoming is that their analysis is based on very few observations. Besides,

⁸In this respect, Brunnermeier (2009) highlights the fact that a reduction in funding liquidity and the subsequent deleveraging process is detrimental for a bank only, if it has to sell assets at fire-sale prices, i.e. if the market liquidity of its assets is low.

⁹Note that the short-term liquidity coverage ratio proposed by the Basel Committee on Banking Supervision is aimed at this effect.

they do not observe the amount lent by an individual bank but only the total amount lent by the syndicate.

In terms of hypothesis number two, the study by Kashyap & Stein (2000) is probably the only point of reference up to now. Their analysis concentrates on the monetary-transmission mechanism and in particular on the bank lending channel. The latter predicts that a monetary contraction, which drains reserves from the banking sector, negatively affects the lending behaviour of banks with less liquid balance sheets – and less access to wholesale funds markets – more than of banks with more liquid assets. The argument is mainly based on the consideration that banks with more liquid assets can easily cover a loss in reserves, while banks with less liquid assets cannot and therefore they have to reduce lending. The findings of the two authors support this argument and indicate the existence of a separate bank lending channel next to the conventional interest-rate channel. Considering hypothesis number two, the question is whether this result only holds for a loss in reserves or also for a sudden reduction in the availability of wholesale funds.

In order to answer this question and to gather further evidence regarding the impact of wholesale dependence on bank lending during a liquidity crisis, we test our two hypotheses empirically in the following part of this chapter. Our analysis involves a number of refinements compared to previous studies, including the fact that we use a large dataset that covers all insured U.S. commercial banks. Furthermore, as we have already outlined above, our analysis is particularly interesting from the perspective of banking regulation since our results have implications for the adequacy of the proposed minimum global liquidity standards.

4.3 Empirical Analysis

Testing our two hypotheses empirically requires that we analyse bank lending behaviour at the individual bank level. In particular, we have to determine how certain bank characteristics, such as wholesale funds dependence, influenced subsequent bank lending. Thus, we use cross-sectional data for our analysis. This micro-econometric approach also helps us to confront the problem of disentangling loan supply from loan demand effects that arises with aggregate loan data.

However, as noted by Ivashina & Scharfstein (2010), using individual bank level data still does not fully eliminate this problem. For instance,

if banks, which are more dependent on wholesale funds, also tend to lend to firms whose loan demand decreases more during a crisis – an example would be investment banks and loans for merger and acquisitions – then finding a negative relation between wholesale dependence and subsequent lending could be the result of loan demand rather than loan supply. Our specification, in turn, which we present in the following section, mitigates this problem by assuming that the marginal effect of wholesale dependence on lending is a function of bank liquidity. If we find evidence that this effect actually varies with liquidity then it is hard to argue that our results are due to changes in loan demand.¹⁰

4.3.1 Variables and Specification

As already indicated above, for our analysis we regress the growth rate of the amount of total loans outstanding in different time periods on a number of indicators at certain time points. For instance, loan growth of bank i in the period from the second quarter of 2007 to the second quarter of 2008, which we denote as $g_{L,i,Crisis1}$, is regressed on balance sheet indicators at the end of the second quarter of 2007. For comparative purposes we also regress the growth rate of bank i in the period from the second quarter of 2008 to the second quarter of 2009, denoted as $g_{L,i,Crisis2}$, on balance sheet indicators at the end of the second quarter of 2008. The main reason for choosing these particular time periods is that the *Crisis1* period contains the beginning of the crisis and the onset of the turmoil in wholesale funding markets while the height of the crisis, i.e. the bankruptcy of Lehman Brothers, and the peak of the liquidity crunch fall into the *Crisis2* period – see also Figure 4.1. The equations we estimate with OLS have the following form:

$$\begin{aligned} g_{L,i,CrisisT} = & \alpha + \beta_1 WD_i + \beta_2 WD_i * \widetilde{LQ}_i + \beta_3 EQ_i \\ & + \beta_4 UC_i + \beta_5 g_{L,i,CrisisT-1} + \beta_6 S_i + u_i \end{aligned} \quad (4.1)$$

with $T = \{1, 2\}$. Note that the regression where the dependent variable is $g_{L,i,Crisis1}$ includes the explanatory variable $g_{L,i,Crisis0}$ on the right hand side. This variable is the growth rate of the amount of total loans outstanding in

¹⁰Notice that one would have to explain why the lending behaviour of banks with the same degree of wholesale dependence, i.e. which face a similar decrease in loan demand, differs.

the time period from the second quarter of 2006 to the second quarter of 2007, i.e. the rate of growth before the crisis. The reason for actually including a lag of the dependent variable is to control for unobserved individual bank characteristics. Another control is bank size S which is calculated as the fraction of the assets of bank i to the total assets of all banks in the sample.

The amount of total loans outstanding, on which our growth rates are based, is given by the sum of real estate loans, agricultural loans, commercial and industrial loans as well as consumer loans – for a detailed description of how we calculate our variables see Table 4.5 in the appendix. Notice that this amount is a stock variable and that ideally we would use data on loan flows to calculate the growth rates used in our regressions. However, since flow data are not available we have to use stocks.¹¹ The main problem with the stock of loans outstanding is that it does not only change because of new loans but also because of loan retirements and drawdowns on existing credit lines. The latter is of particular concern given the evidence for a run by borrowers. In order to control for an increase in the amount of total loans outstanding during the crisis due to this run we include the ratio of unused commitments to total assets, denoted by UC , in all regressions. The coefficient sign of this variable is expected to be positive, i.e. banks with a higher amount of unused commitments are more likely to increase subsequent lending due to a run by borrowers.

Another control variable is EQ which is the ratio of equity capital to total assets. The reason why we include this variable in our specification is the following. Assume that a shock reduces both equity capital and the value of risk weighted assets of a bank by an amount ϵ and the bank is unable to raise any new equity. The question in this case is by how much does the bank have to decrease risk weighted assets in order to maintain the equity to risk weighted asset ratio from before the shock. When equity is given by E_t before the shock and risk weighted assets are given by A_t then we are searching for the value of x that fulfils the condition $(E_t - \epsilon)/(A_t - \epsilon - x) = E_t/A_t$. Also note that the value of x can be interpreted as the amount the bank has to deleverage. It is straightforward to show that $x = \epsilon(A_t - E_t)/E_t$ and that $\partial x/\partial E_t < 0$. In other words, a bank with a bigger capital buffer before the shock will have to decrease risk weighted assets, including loans, less than

¹¹In this respect Bernanke & Lown (1991) argue that if loan maturities are rather long then the real value of new loans is best approximated with the nominal growth rate of loans outstanding.

a bank with a smaller buffer. Therefore including EQ in our regressions controls for another channel besides the liquidity crunch that could have affected lending during the crisis. We expect the coefficient of the equity variable to have a positive sign.

Let us now turn to the main variables of interest, namely wholesale dependence WD and liquidity LQ . Instead of using the former, many authors work with wholesale independence which they define as the ratio of total deposits to total assets. Unfortunately, though, this definition does not take into account that a certain fraction of total deposits, such as brokered deposits, are wholesale funds. In calculating an indicator for wholesale dependence we therefore exclude deposits from domestic and foreign banks and measure dependence as the ratio of total liabilities minus equity and adjusted deposits to total assets. In the same way, finding an adequate indicator for the liquidity of a bank's balance sheet is not a trivial task. Kashyap & Stein (2000), for instance, use the ratio of total security holdings to total assets. However, this measure does not account for the fact that during the recent financial crisis the market liquidity of many securities declined sharply, implying that banks could sell these securities only with a huge discount. For this reason our liquidity indicator only includes security holdings which were likely to be highly liquid even during the crisis.¹²

Our second hypothesis implies that equation (4.1) should not only include wholesale dependence as explanatory variable but also an interaction of the latter with liquidity. If we would interact WD and LQ , the marginal effect of wholesale dependence on subsequent lending would be given by $\beta_1 + \beta_2 LQ_i$. The coefficient β_1 would then represent the marginal effect of wholesale dependence on lending if $LQ_i = 0$, i.e. if bank i does not hold any liquid assets which is a rather extreme case. Therefore we reformulate our specification in the following way. Instead of using LQ_i in the interaction term we use $\widetilde{LQ}_i = LQ_i - Q_{z,LQ}$, where $Q_{z,LQ}$ is the z -quantile of the LQ series – later on we limit our analysis to the three quartiles, i.e. $z = \{0.25, 0.5, 0.75\}$. Formulating our model in this way the coefficient of WD is immediately interpretable and we also obtain standard errors for the marginal effect of wholesale dependence at meaningful levels of LQ . Given

¹²With regard to liquidity, Kashyap & Stein (2000) indicate the following problem. If a bank operates in an area that offers few good lending opportunities then the bank will hold more liquid assets rather than making bad loans. In this case liquidity would be an endogenous variable. However, since we analyse subsequent lending behaviour, our explanatory variables are predetermined.

this particular specification, our two hypotheses imply that the coefficient on WD is negative (Hypothesis #1) while the coefficient of the interaction term is positive, i.e. for higher levels of liquidity, the negative effect of wholesale dependence should become smaller in absolute terms (Hypothesis #2).

4.3.2 Data

The data, on which our analysis and the results presented in the following section are based, have been taken from the Consolidated Reports of Condition and Income, the so-called Call Reports, in which federally insured banks in the United States disclose information from their income statement and balance sheet on a quarterly basis.¹³ The data are submitted to the Federal Deposit Insurance Corporation (FDIC) and they cover banks which are regulated by the Federal Reserve System, the FDIC and the Comptroller of the Currency.

The main advantage of Call Report data is that they provide a comprehensive dataset that covers most of the banks operating in the United States. The main difficulties with respect to this huge data pool are that small banks do not report as much data as large banks and that the variables and their measurement change over time in order to account for various developments in the banking sector – the second point implies that it can be difficult to form consistent time series. Notice, however, that neither of these two issues affects our analysis. Furthermore, as we have already pointed out above, the balance sheet data obtained from the Call Reports are stock variables.

Another issue that deserves closer attention is the distinction between consolidated and domestic data series. The former, identified with the code RCFD, not only include domestic but also foreign branches of a bank while the latter, with the identifier RCON, only include domestic branches. The choice which series a researcher should use depends on the type of analysis. Kashyap & Stein (2000), for instance, analyse changes in bank lending behaviour due to monetary policy. Therefore they should use domestic data only because the lending decisions of foreign branches are probably less dependent on changes in U.S. monetary policy. In our case, in turn, we are

¹³The data we use can be downloaded from the website of the Federal Reserve Bank of Chicago (<http://www.chicagofed.org/webpages/banking/index.cfm>). A good introduction to Call Report bank data by den Haan et. al. can be downloaded from <http://www.csulb.edu/~gyamashi/manual.pdf>.

interested in changes of bank lending behaviour due to individual bank characteristics. Thus, we opt for using consolidated data.¹⁴

Nevertheless, we restrict our sample of banks in order to ensure a certain degree of homogeneity such that the influence of unobserved individual bank characteristics is reduced. The restrictions to which we subject our dataset are given in Table 4.4 in the appendix. As one can see from this table, we concentrate on commercial banks – e.g. we exclude savings banks or credit unions – which are insured by the FDIC and which operate within the fifty states of the U.S. plus the District of Columbia. The reason, why we exclude banks operating in exotic places such as the U.S. Virgin Islands, for instance, is that these are different business environments.

In terms of outliers, our criterion of exclusion is based on the distance of a particular observation from the 0.25- and the 0.75-quartile of the sample respectively. For instance, an observation with an equity to asset ratio of 90% constitutes a very unusual observation that is far beyond the threshold of the 0.75-quartile plus three times the interquartile range, wherefore we exclude such an observation.

Finally, we restrict our sample to banks with a total loan to asset ratio of at least 45%. On the one hand, this step is likely to reduce the heterogeneity among the banks we analyse even further. On the other hand, since we are interested in changes of bank lending behaviour, this restriction guarantees that our sample only includes banks with a sizeable lending business. All in all, the banks in our two samples – one for the *Crisis1* and another for the *Crisis2* period – account for approximately three quarters of the sum of total loans outstanding before the implementation of this restriction.

4.3.3 Estimation Results

The main results of our analysis are summarised in Table 4.2.¹⁵ Looking at column number one, where the dependent variable is $g_{L,i,Crisis1}$, we see that the coefficients of wholesale dependence, given that $LQ = Q_{.25,LQ}$, and of the

¹⁴However, in case that RCFD data are not available for a particular series we have to use RCON data. See also Table 4.5.

¹⁵The values in parenthesis are t-statistics. The stars indicate the level of significance, i.e. *** stand for the 1%-level of significance, ** for the 5%-level of significance and * denotes the 10%-level of significance. Also note that in our initial regressions we found evidence of heteroskedastic residuals. For this reason we estimated White Heteroskedasticity-Consistent standard errors.

interaction with liquidity are both significantly different from zero. Thus, the marginal effect of wholesale dependence is actually related to liquidity and, more importantly, a higher wholesale dependence decreases subsequent lending growth. For higher values of liquidity, however, the effect is not significantly different from zero.

With regard to the other explanatory variables we observe that except for equity all variables have a statistically significant coefficient. *Ceteris paribus*, these results indicate that banks with more unused commitments before the crisis had a higher subsequent growth rate of lending, which we interpret as evidence for the run of borrowers. Furthermore, banks with a higher growth rate of lending in the pre-crisis period also had a higher growth rate in the first phase of the crisis. And finally, larger banks witnessed a smaller growth rate during the crisis. Also notice that our goodness-of-fit is not too bad given that we work with cross-sectional data.

For the *Crisis2* period, in turn, we basically get the same results – see column two in Table 4.2. The main difference is that now the negative effect of wholesale dependence is stronger and significant both at $LQ = Q_{.25,LQ}$ and $LQ = Q_{.5,LQ}$. Besides, the coefficient of equity is now significantly different from zero and it has the expected positive sign. Unused commitments and size, though, are less significant and the goodness-of-fit decreases slightly.¹⁶

These results support our first hypothesis, namely that more wholesale dependent banks decreased loan supply more than less dependent banks. With regard to our second hypothesis, we observe that in both periods the coefficient of the interaction between wholesale dependence and liquidity is statistically significant and positive. In other words, for higher levels of liquidity the negative effect of wholesale dependence on subsequent lending growth becomes smaller in absolute terms. Thus, our results also favour our second hypothesis that wholesale dependent banks with a more liquid balance sheet decreased loan supply less than banks with less liquid balance sheets.

Moreover, our results are not sensitive to the indicator used for wholesale dependence and liquidity respectively.¹⁷ Even if we run our regressions with

¹⁶Although the distributions of our residuals are not exactly normal but leptokurtic, they have zero mean and due to the large number of observations our statistical tests still seem to be valid.

¹⁷The alternative indicators we have used for our sensitivity analysis include the wholesale dependence indicator that does not account for bank deposits and the liquidity measure proposed by Kashyap & Stein (2000). See also section 4.3.1 for more information.

	(1)	(2)
Dependent Variable	$g_{L,Crisis1}$	$g_{L,Crisis2}$
<i>const.</i>	0.0211 (2.7744)***	0.0037 (0.5002)
<i>WD</i> (if $LQ = Q_{.25,LQ}$)	-0.0932 (-3.8786)***	-0.1536 (-7.3623)***
<i>WD</i> (if $LQ = Q_{.5,LQ}$)	-0.0374 (-1.5773)	-0.1014 (-4.8584)***
<i>WD</i> (if $LQ = Q_{.75,LQ}$)	0.0319 (1.0767)	-0.0322 (-1.1439)
$WD * \widetilde{LQ}$	1.4294 (5.0837)***	1.2786 (4.8830)***
<i>EQ</i>	-0.0372 (-0.6092)	0.1447 (2.3876)**
<i>UC</i>	0.2732 (11.8276)***	0.0472 (1.9897)**
$g_{L,CrisisT-1}$	0.3363 (21.3672)***	0.2702 (18.3567)***
<i>S</i>	-1.2875 (-3.7025)***	-2.6420 (-1.8503)*
<i>Obs.</i>	5,489	5,389
<i>adj. R²</i>	0.17	0.11

Table 4.2: Estimation Results

	95% Confidence Interval	
	<i>Crisis1</i>	<i>Crisis2</i>
if $LQ = Q_{.25,LQ}$	$[-0.14, -0.05]$	$[-0.19, -0.11]$
if $LQ = Q_{.5,LQ}$	$[-0.08, 0.01]$	$[-0.14, -0.06]$
if $LQ = Q_{.75,LQ}$	$[-0.02, 0.09]$	$[-0.09, 0.02]$

Table 4.3: Confidence Intervals of WD Coefficient

samples including banks with a total loans to asset ratio of below 45% the results do not change dramatically and essentially stay the same.

The main question that arises from these results is whether they are also meaningful from an economic perspective. In order to answer this question, consider a bank whose balance sheet was rather illiquid at the end of the second quarter of 2007, i.e. liquidity was equal to the 0.25-quartile of the LQ series. A marginal increase in wholesale funds dependence then implied that the growth rate of the amount of total loans outstanding in the *Crisis1* period decreased by five to fourteen percentage points – in the *Crisis2* period the estimated reduction would have been even larger.¹⁸ On the other hand, a bank with a very liquid balance sheet, i.e. with liquidity equal to the 0.75-quartile of the LQ series, would have witnessed a change in the growth of the amount of loans outstanding of minus two to plus nine percentage points due to a marginal increase in wholesale dependence. This suggests that a combination of wholesale dependence and illiquid asset holdings can in fact have a detrimental effect on the lending behaviour of individual banks during a liquidity crunch.

Finally, with regard to banking regulation, the empirical evidence gathered from our analysis supports the introduction of global minimum liquidity standards. In particular, our results concerning the relation between wholesale dependence and liquidity indicate that a short-term liquidity coverage ratio that forces banks to hold more liquid assets is likely to mitigate the effect of a disruption in wholesale funding access, at least in terms of lending. The significant negative effect of wholesale dependence on subsequent lending further suggests that curtailing the use of this type of funds and fostering the use of more stable sources of funds could be warranted, although we have to emphasise that this result is related to the lending behaviour during

¹⁸See Table 4.3 which contains the 95% confidence intervals for the coefficient of wholesale dependence for each quartile of LQ and for each sample period.

a liquidity crisis and not lending in general. Besides, our findings in terms of equity capital holdings reinforce the view that banks should hold enough (high-quality) capital. In other words, banks should have an incentive to operate with a reasonable leverage only.

4.4 Conclusion

During the financial crisis that started in 2007 banks dependent on wholesale funds had increasing difficulties to roll over their short-term debts because wholesale financiers withdrew their funds. There is widespread agreement that this liquidity crunch resulted in a vicious deleveraging cycle that amplified the crisis and helped to propagate it. In the course of this cycle banks had to shrink the asset side of their balance sheets in order to meet obligations due, implying that they had to sell securities and/or stop lending. Against this background, this paper tests the hypothesis that more wholesale dependent banks reduced subsequent lending more than less dependent banks. In addition, we analyse whether wholesale dependent banks with a more liquid balance sheet decreased lending less than banks with less liquid assets. The results of our empirical analysis, which covers the time period from the second quarter of 2007 to the second quarter of 2008 and the period from the second quarter of 2008 to the second quarter of 2009, indicate that neither of these two hypotheses can be rejected. Furthermore, our results are not only insensitive to changes in the measurement of wholesale dependence and liquidity respectively but they are also significant from an economic perspective. The empirical evidence therefore supports the introduction of global minimum liquidity standards as proposed by the Basel Committee on Banking Supervision and reinforces the proposal to increase the quantity (and quality) of bank equity capital.

4.5 Appendix

Call Report Code	Name	Values
rssd9048	Charter Type	200 (Commercial Bank)
rssd9331	Entity Type	1 (Commercial Bank)
rssd9424	Primary Insurer	1, 2, 6 or 7 (FDIC)
rssd9210	Physical State	1-56 (the fifty states of the U.S. plus DC)

Table 4.4: Determination of Banking Sample

Variable	Name	Formula & Call Report Codes
L	Total Loans Outstanding	$\text{rcfd1410} + \text{rcfd1590} + \text{rcfd1766} + \text{rcfd1975}$
$gL, Crisis X$	Change in Total Loans Outstanding	$(LC_{Crisis X} - LC_{Crisis X-1}) / LC_{Crisis X-1}$
WD	Wholesale Dependence	$1 - [\text{rcfd2200} \cdot (\text{rconB551} + \text{rconB552}) \cdot (\text{rcon2213} + \text{rcon2236}) + \text{rcfd3210}] / \text{rcfd2170}$
LQ	Liquidity	$(\text{rcfd0010} + \text{rcfdB987} + \text{rcfdB989} + \text{rcfd0211} + \text{rcfd1286} + \text{rcfd8496} + \text{rcfd8498} + \text{rcon1737} + \text{rcon1739} + \text{rcon1742} + \text{rcon1744} + \text{rconA510}) / \text{rcfd2170}$
EQ	Equity Ratio	$\text{rcfd3210} / \text{rcfd2170}$
UC	Unused Commitments Ratio	$(\text{rcfd3814} + \text{rcfd3815} + \text{rcfd3816} + \text{rcfd3818}) / \text{rcfd2170}$
S	Size	$\text{rcfd2170} / \sum_i \text{rcfd2170}$

Table 4.5: Variable Definitions

Bibliography

- Akhtar, M. A. (1993), ‘Causes and consequences of the 1989-1992 credit slowdown: Overview and perspectives’, *Federal Reserve Bank of New York Quarterly Review* **18(4)**, 1–23.
- Allain, L. & Oulidi, N. (2009), ‘Credit market in Morocco: A disequilibrium approach’, *International Monetary Fund Working Paper* **53**.
- Baltensperger, E. (1978), ‘Credit rationing. Issues and questions’, *Journal of Money, Credit, and Banking* **10(2)**, 170–183.
- Barajas, A. & Steiner, R. (2002), ‘Credit stagnation in Latin America’, *International Monetary Fund Working Paper* **53**.
- Berger, A. N. & Udell, G. F. (1992), ‘Some evidence on the empirical significance of credit rationing’, *Journal of Political Economy* **100(5)**, 1047–1077.
- Berger, A. N. & Udell, G. F. (1994), ‘Did risk-based capital allocate bank credit and cause a “credit crunch” in the United States?’, *Journal of Money, Credit, and Banking* **26(3)**, 585–628.
- Bernanke, B. S. & Lown, C. S. (1991), ‘The credit crunch’, *Brookings Papers on Economic Activity* **2**, 205–247.
- Blanchard, O. J. (2008), ‘The crisis: Basic mechanisms, and appropriate policies’, *MIT Economics Department Working Paper Series* **09-01**.
- Bowden, R. J. (1978), ‘Specification, estimation and inference for models of markets in disequilibrium’, *International Economic Review* **19(3)**, 711–726.

- Brinkmann, E. J. & Horvitz, P. M. (1995), ‘Risk-based capital standards and the credit crunch’, *Journal of Money, Credit, and Banking* **27(3)**, 848–863.
- Browne, F. X. (1987), ‘Sluggish quantity adjustment in a non-clearing market - a disequilibrium econometric application to the loan market’, *Journal of Applied Econometrics* **2(4)**, 335–349.
- Brunnermeier, M. K. (2009), ‘Deciphering the liquidity and credit crunch 2007-2008’, *Journal of Economic Perspectives* **23(1)**, 77–100.
- Brunnermeier, M. K. & Pedersen, L. H. (2009), ‘Market liquidity and funding liquidity’, *Review of Financial Studies* **22(6)**, 2201–2238.
- Calomiris, C. W. (1999), ‘Building and incentive-compatible safety net’, *Journal of Banking and Finance* **23(10)**, 1499–1519.
- Chari, V. V., Christiano, L. & Kehoe, P. J. (2008), ‘Facts and myths about the financial crisis of 2008’, *Federal Reserve Bank of Minneapolis Working Paper* **666**.
- Contessi, S. & Francis, J. L. (2009), ‘U.S. commercial bank lending through 2008:Q4: New evidence from gross credit flows’, *Federal Reserve Bank of St. Louis Working Paper Series* **11**.
- Diamond, D. W. & Dybvig, P. H. (1983), ‘Bank runs, deposit insurance, and liquidity’, *Journal of Political Economy* **91(3)**, 401–419.
- Drehmann, M. & Nikolaou, K. (2009), ‘Funding liquidity risk. Definition and measurement’, *European Central Bank Working Paper Series* **1024**.
- Fair, R. C. & Jaffee, D. M. (1972), ‘Methods of estimation for markets in disequilibrium’, *Econometrica* **40(3)**, 497–514.
- Feldman, R. J. & Schmidt, J. (2001), ‘Increased use of uninsured deposits’, *fedgazette* **March**.
- Fermanian, J.-D. & Salanie, B. (2004), ‘A nonparametric simulated maximum likelihood estimation method’, *Econometric Theory* **20(4)**, 701–734.
- Friedman, B. M. (1991), ‘Comment on “The Credit Crunch” by Ben S. Bernanke and Cara Lown’, *Brookings Papers on Economic Activity* **2**, 240–244.

- Goldfeld, S. M. & Quandt, R. E. (1981), 'Single-market disequilibrium models : Estimating and testing', *Journal of Econometrics* **16(1)**, 157.
- Gosh, S. R. & Gosh, A. R. (1999), 'East Asia in the aftermath: Was there a crunch?', *International Monetary Fund Working Paper* **38**.
- Gourieroux, C., Monfort, A. & Trognon, A. (1985), 'A general approach to serial correlation', *Econometric Theory* **1(3)**, 315–340.
- Hajivassiliou, V. A. (1986), 'Two misspecification tests for the simple switching regressions disequilibrium model', *Economics Letters* **22(4)**, 343–348.
- Harris, D. G. (1974), 'Credit rationing at commercial banks: Some empirical evidence', *Journal of Money, Credit, and Banking* **6(2)**, 227–240.
- Huang, R. & Ratnovski, L. (2011), 'The dark side of bank wholesale funding', *Journal of Financial Intermediation* **20(2)**, 248–263.
- Hwang, H. (1980), 'A test of a disequilibrium model', *Journal of Econometrics* **12(3)**, 319–333.
- Ikhide, S. (2003), 'Was there a credit crunch in Namibia between 1996-2000?', *Journal of Applied Economics* **6(2)**, 269–290.
- Ito, T. & Ueda, K. (1981), 'Tests of the equilibrium hypothesis in disequilibrium econometrics: An international comparison of credit rationing', *International Economic Review* **22(3)**, 691–708.
- Ivashina, V. & Scharfstein, D. (2010), 'Bank lending during the financial crisis of 2008', *Journal of Financial Economics* **97(3)**, 319–338.
- Jaffee, D. M. & Modigliani, F. (1969), 'A theory and test of credit rationing', *American Economic Review* **59(5)**, 850–872.
- Kashyap, A. K. & Stein, J. C. (2000), 'What do a million observations on banks say about the transmission of monetary policy?', *American Economic Review* **90(3)**, 407–428.
- Kim, H. E. (1999), 'Was the credit channel a key monetary transmission mechanism following the recent financial crisis in the Republic of Korea?', *World Bank Policy Research Working Paper* **2103**.

- King, S. R. (1986), ‘Monetary transmission: Through bank loans or bank liabilities?’, *Journal of Money, Credit, and Banking* **18(3)**, 290–303.
- Krishnamurthy, A. (2010), ‘How debt markets have malfunctioned in the crisis’, *Journal of Economic Perspectives* **24(1)**, 3–28.
- Kugler, P. (1987), ‘Credit rationing and the adjustment of the loan rate: An empirical investigation’, *Journal of Macroeconomics* **9(4)**, 505–525.
- Laffont, J.-J. & Garcia, R. (1977), ‘Disequilibrium econometrics for business loans’, *Econometrica* **45(5)**, 1187–1204.
- Laroque, G. & Salanie, B. (1993), ‘Simulation-based estimation of models with lagged latent variables’, *Journal of Applied Econometrics* **8(S1)**, 119–133.
- Laroque, G. & Salanie, B. (1994), ‘Estimating the canonical disequilibrium model: Asymptotic theory and finite sample properties’, *Journal of Econometrics* **62(2)**, 165–210.
- Lee, L.-F. (1997), ‘A smooth likelihood simulator for dynamic disequilibrium models’, *Journal of Econometrics* **78(2)**, 257–294.
- Lown, C. S. & Morgan, D. P. (2006), ‘The credit cycle and the business cycle: New findings using the loan officer opinion survey’, *Journal of Money, Credit, and Banking* **38(6)**, 1575–1597.
- Maddala, G. S. (1983), *Limited-dependent and Qualitative variables in econometrics*, Cambridge University Press.
- Maddala, G. S. & Nelson, F. D. (1974), ‘Maximum likelihood methods for models of markets in disequilibrium’, *Econometrica* **42(6)**, 1013–1030.
- Mayer, W. J. (1989), ‘Estimating disequilibrium models with limited a priori price-adjustment information’, *Journal of Econometrics* **41(3)**, 303–320.
- Melitz, J. & Pardue, M. (1973), ‘The demand and supply of commercial bank loans’, *Journal of Money, Credit, and Banking* **5(2)**, 669–692.
- Mishkin, F. S. (2004), *The Economics of Money, Banking, and Financial Markets*, Addison-Wesley.

- Nehls, H. & Schmidt, T. (2003), ‘Credit crunch in Germany?’, *RWI Discussion Paper* **6**.
- Pazarbasioglu, C. (1997), ‘A credit crunch? Finland in the aftermath of the banking crisis’, *International Monetary Fund Staff Paper* **44(3)**, 315–327.
- Peek, J. & Rosengren, E. (1995), ‘The capital crunch: Neither a borrower nor a lender be’, *Journal of Money, Credit, and Banking* **27(3)**, 625–638.
- Poghosyan, T. (2010), ‘Slowdown of credit flows in Jordan in the wake of the global financial crisis: Supply or demand driven?’, *International Monetary Fund Working Paper* **256**.
- Quandt, R. E. (1988), *The Econometrics of Disequilibrium*, Blackwell.
- Raddatz, C. (2010), ‘When the rivers run dry: Liquidity and the use of wholesale funds in the transmission of the U.S. subprime crisis’, *World Bank Policy Research Working Paper* **5203**.
- Rudebusch, G. D. (1987), *The Estimation of Macroeconomic Disequilibrium Models with Regime Classification Information*, Springer-Verlag.
- Sealey, C. W. (1979), ‘Credit rationing in the commercial loan market: Estimates of a structural model under conditions of disequilibrium’, *Journal of Finance* **34(3)**, 689–702.
- Shin, H. S. (2009), ‘Reflections on Northern Rock: The bank run that heralded the global financial crisis’, *Journal of Economic Perspectives* **23(1)**, 101–119.
- Srivastava, V. K. & Rao, B. B. (1990), *The Econometrics of Disequilibrium Models*, Greenwood Press.
- Stiglitz, J. E. & Weiss, A. (1981), ‘Credit rationing in markets with imperfect information’, *American Economic Review* **71(3)**, 393–410.
- Wojnilower, A. M. (1994), *The New Palgrave Dictionary of Money and Finance*, Palgrave Macmillan, pp. 525–527.

Abstract

This thesis deals with the estimation of disequilibrium econometric models as well as the implications of the financial crisis that started in 2007 on bank lending. While the first article is solely concerned with disequilibrium models, the third article mainly focuses on the implications of the financial crisis on bank lending behaviour. The second article, in contrast, brings these two research topics together.

Many markets, like the labour or the credit market, differ in some important respects from the standard market framework in economics. Disequilibrium models account for this peculiarity in the sense that they do not assume that markets necessarily clear. Two major issues in the estimation of these models are the identification of the most appropriate model specification and the provision for dynamic model features. In the first article we show by means of Monte Carlo sampling experiments that although the degree of misspecification of the canonical disequilibrium models depends on the size of a particular model parameter, researchers should pay careful attention to the issue of specification. Our results further suggest that the estimators of the static canonical models, though they are consistent under certain conditions, are not reliable in dynamic settings which calls for the use of tractable dynamic models.

In the second article we apply the disequilibrium framework to U.S. loan market data. According to conventional wisdom the recent financial crisis entailed a so-called credit crunch, i.e. banks cut back on lending sharply essentially rationing the amount of credit. However, empirical evidence regarding credit rationing in general and the credit crunch hypothesis in particular is ambiguous. Using an alternative excess credit demand indicator – which seems to be superior to previous indicators used – within the disequilibrium framework, we analyse whether U.S. commercial banks have increased non-price credit rationing sharply in the market for commercial and industrial

loans during the crisis. Our findings suggest that although excess demand for credit indeed increased, the extent of rationing is negligible, implying that its real effects are moderate at most.

Finally, the financial crisis witnessed the emergence of a new form of bank run. In this modern form, also referred to as liquidity crunch, it is not depositors that run but wholesale financiers. The consequences, however, are similar to a traditional run, that is banks have to shrink their balance sheets. In the third article we show that banks more vulnerable to the run, i.e. more wholesale dependent banks, decreased loan supply more than less vulnerable banks and that the amount of liquid asset holdings affects lending behaviour positively during a liquidity crunch. Thus, our results support the introduction of liquidity standards as proposed by the Basel Committee on Banking Supervision.

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit der Schätzung von sogenannten Disequilibrium Modellen sowie mit den Auswirkungen der Finanzkrise von 2007 auf die Vergabe von Bankkrediten. Während sich der erste Artikel ausschließlich mit Disequilibrium Modellen befasst, konzentriert sich der dritte Artikel hauptsächlich auf die oben genannten Auswirkungen der Finanzkrise. Im Gegensatz dazu werden diese beiden Forschungsschwerpunkte im zweiten Artikel zusammengeführt.

Viele real existierende Märkte, wie etwa der Arbeits- oder der Kreditmarkt, unterscheiden sich grundlegend von der in der ökonomischen Theorie vorherrschenden Modellvorstellung eines Marktes. Disequilibrium Modelle tragen dieser Diskrepanz unter anderem dadurch Rechnung indem ihnen die Annahme zugrunde liegt, dass Märkte nicht notwendigerweise geräumt werden. Zwei wesentliche Punkte bei der Schätzung von derartigen Modellen sind das Finden der geeignetsten Modellspezifikation sowie die Berücksichtigung von dynamischen Modellkomponenten. Im ersten Artikel dieser Dissertation zeigen wir anhand von Monte Carlo Experimenten, dass obwohl der Grad der Misspezifikation der kanonischen Disequilibrium Modelle von einem Modell-Parameter abhängt, Forscher sehr vorsichtig bei der Auswahl der Modellspezifikation agieren sollten. Zudem legen unsere Resultate nahe, dass die Verwendung von Schätzern, die auf den kanonischen, d.h. statischen, Disequilibrium Modellen basieren, keine zuverlässigen Resultate liefert wenn der zugrundeliegende datengenerierende Prozess dynamisch ist. Dieser Umstand impliziert, dass in der Praxis auf Schätzer zurückgegriffen werden sollte, die von dynamischen Modellen abgeleitet wurden.

Im zweiten Artikel wenden wir schließlich Disequilibrium Modelle auf Daten des U.S. Kreditmarktes an. Laut gängiger Meinung hatte die Finanzkrise von 2007 einen sogenannten Credit Crunch zur Folge, d.h. Banken reduzierten ihr Kreditangebot drastisch was einer Rationierung der Kredit-

menge gleichkam. Was jedoch die empirische Beweislage mit Hinblick auf Kreditrationierung im Allgemeinen und die Credit Crunch Hypothese im Speziellen betrifft, so kann man von keiner eindeutigen Richtung sprechen. Vor diesem Hintergrund analysieren wir ob Geschäftsbanken in den Vereinigten Staaten während der Krise tatsächlich das Angebot für Industriekredite drastisch rationierten. Dabei wenden wir im Rahmen eines Disequilibrium Modells einen alternativen Indikator für den Nachfrageüberschuss bei Industriekrediten an, der besser geeignet zu sein scheint als bisherige Indikatoren. Unsere Resultate zeigen, dass obwohl es einen Anstieg bei der Rationierung von Kredit gegeben haben dürfte, dieser vernachlässigbar klein war, was wiederum impliziert, dass die realen Auswirkungen dieses Anstiegs bestenfalls moderat waren.

Schließlich beschäftigt sich der dritte Artikel mit einer neuen Form des Bank Runs. Bei dieser neuen Form, die auch Liquidity Crunch genannt wird, sind es nicht Privat- bzw. Kleinkunden die ihre Einlagen in großer Zahl zurückfordern sondern Großinvestoren, wie etwa andere Banken oder Fonds, die einer Bank Finanzmittel entziehen. Die Auswirkungen sind ähnlich wie bei einem herkömmlichen Bank Run, nämlich dass die betroffene Bank ihre Bilanz schrumpfen muss. In diesem Zusammenhang zeigt unsere Untersuchung, dass Banken, die anfälliger für diese neue Form des Bank Run waren, d.h. Banken die mehr auf Großinvestoren angewiesen waren, ihr Kreditangebot stärker reduzierten als weniger anfällige Banken. Zudem weisen unsere Ergebnisse darauf hin, dass dieser negative Effekt für Banken, die über ein größeres Polster aus liquiden Mitteln verfügten, schwächer war. Somit stützen unsere Resultate die Einführung von Liquiditätsstandards wie sie etwa vom Baseler Komitee für Bankenaufsicht vorgeschlagen wurden.

Curriculum Vitae

Personal Data

Name: Thomas Hrdina
Address: Kaiserstraße 83/13, A-1070 Wien
E-Mail: thomas.hrdina@univie.ac.at
Date of Birth: September 8, 1980

Education

since 03/2008 **PhD Programme in Economics**
University of Vienna (graduation expected autumn 2011)
2004–2008 **Master’s Programme in Economics**
University of Vienna (diploma with distinction)
2003–2004 **BA(Hons) Management, Business & Administration**
New College Durham, England (graduation with first-class honours)
2000–2003 **International Business Administration**
European Management Academy, Vienna (diploma with distinction)

Academic Positions

01/2009–06/2011 **Department of Economics**, University of Vienna
(Research Assistant)
10/2007–01/2008 **Department of Finance**, University of Vienna
(Teaching assistant)
10/2006–01/2007 **Department of Finance**, University of Vienna
(Teaching assistant)

Professional Experience

07/2008–09/2008 **Raiffeisen Zentralbank Österreich AG**, Vienna
(Secondary Loan Trading & Distribution Dept.)
07/2007–08/2007 **Raiffeisen Zentralbank Österreich AG**, Vienna
(Economics & Financial Market Research Dept.)
02/2007–06/2007 **European Parliament**, Brussels
(Assistant in the office of Dr. Hannes Swoboda, MEP)

before 10/2006 **Raiffeisen Zentralbank Österreich AG**, Vienna
 Oesterreichische Kontrollbank AG, Vienna
 Funiversal Travel Agency S.A., Madrid
 Österreichische Volksbanken AG, Vienna
 Invesco Bank Österreich AG, Vienna

Teaching Experience

Winter 2010 **Introductory Macroeconomics**, University of Vienna
 (course for undergraduate business administration students)

 Summer 2010 **Introductory Macroeconomics**, University of Vienna
 (course for undergraduate business administration students)

 Winter 2009 **Introductory Macroeconomics**, University of Vienna
 (course for undergraduate business administration students)

 Summer 2009 **Introductory Macroeconomics**, University of Vienna
 (course for undergraduate business administration students)

Research Interests

Economics of Banking, Macroeconomics and Applied Econometrics